DEX-0245 BC

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 March 2002 (07.03.2002)

PCT

(10) International Publication Number
WO 02/18424 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/US01/27093

(22) International Filing Date: 31 August 2001 (31.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/654,935 1 September 2000 (01.09.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/654,935 (CIP)
Filed on 1 September 2000 (01.09.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHOU, Ping** [US/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendahl Lane, Cupertino, CA 95014 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **WANG, Dunrui** [CN/US]; 12252 Pepper Tree Lane, Poway, CA 92064 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [US/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). **WEHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P., C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/18424 A2

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, circulating soluble factors, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-245. The polypeptides sequences are designated SEQ ID NO: 246-490. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is unknown or any of the four bases.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-245 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1-245. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-245 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-245. The sequence information can be a segment of any one of SEQ ID NO: 1-245 that uniquely identifies or represents the sequence information of SEQ ID NO: 1-245.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information are provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization

probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-245 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-245 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-245; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1-245; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-245. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-245; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in SEQ ID NO: 246-490; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO: 1-245; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provide methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2); for which they have a signature region (as set forth in Table 3); or for which they have homology to a gene family (as set forth in Table 4). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady

and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30

nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NO: 1-245.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-245. The sequence information can be a segment of any one of SEQ ID NO: 1-245 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO: 1-245. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 500 amino acids, more preferably less than 200 amino acids more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include an initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2): 134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-245; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 246-490; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO: 246-490. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-245; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing as SEQ ID NO: 246-490; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 246-490. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding,

extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-245 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-245 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1-245 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-245, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to) any one of the polynucleotides of the invention are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1-245, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1-245 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO: 1-245, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-245, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-245 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-245 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE NUCLEIC ACIDS

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-245, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO: 246-490 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-245 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-245), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-245). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an mRNA of SEQ ID NO: 1-245 (see, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742). Alternatively, polynucleotides of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 246-490 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-245 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-245 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 246-490 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 246-490 or the corresponding full length or mature protein; and "substantial equivalents" thereof (*e.g.*, with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 246-490.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 246-490.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.
In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.,* cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.
Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular

endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus,

rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic

composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population.

Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostasis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention

may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide,

Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen

recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such

transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding

molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or

disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or

differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or

elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified

nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of

administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic

factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

5

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous,
10 intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

15 Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery
20 system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and
25 an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

30

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into
35 preparations which can be used pharmaceutically. These pharmaceutical compositions may be

manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers

enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with

an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well

known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable

lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions

may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which

modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the

population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the

invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as the amino acid sequences shown in SEQ ID NO: 246-490, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte

Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

4.13.1 POLYCLONAL ANTIBODIES

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the

target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5

4.13.2 MONOCLONAL ANTIBODIES

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain
10 gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those
15 described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.
The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion
20 protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually
25 transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for
30 the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which
35 can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego,

California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.
The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin

polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

4.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

4.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal

antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al,(Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

4.13.5 F_{ab} FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

4.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the

binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to

stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular

defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

4.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

4.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

4.13.9 IMMUNOCONJUGATES

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon

a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO: 1-245 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO: 1-245 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage

means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see

Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard,

T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the

invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO: 1-245, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the

invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems.

Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO: 1-245. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from any of the nucleotide sequences SEQ ID NO: 1-245 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed

(Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/μl) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-$MeIm_7$), is then added to a final concentration of 10 mM 1-$MeIm_7$. The single-stranded DNA solution is then dispensed into CovaLink NH strips (75 μl/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-$MeIm_7$, is made fresh and 25 μl added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *CviJI*, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *CviJI* normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of

this enzyme (*CviJI***), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *CviJI*** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *CviJI*** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 μg instead of 2-5 μg); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic

strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5. EXAMPLES

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences.

5.2 EXAMPLE 2

Assemblage of Novel Nucleic Acids

The nucleic acids of the present invention, were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST

sequences, dbEST, gb pri, UniGene, and exons from public domain genomic sequences predicated by GenScan) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Further, inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), full-length gene sequences and their corresponding protein sequences were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTXY algorithm against Genbank (i.e., dbEST, gb pri, UniGene, and Genpept). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). In some cases RACE (Rapid Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction. The full-length nucleotide sequences are shown in the Sequence Listing as SEQ ID NO: 1-245. The corresponding polypeptide sequences are SEQ ID NO: 246-490.

Table 1 shows the various tissue sources of SEQ ID NO: 1-245.

The nearest neighbor results for polypeptides encoded by SEQ ID NO: 1-245 (i.e. SEQ ID NO: 246-490) were obtained by a BLASTP (version 2.0a1 19MP-WashU) search against Genpept release 124 using BLAST algorithm. The nearest neighbor result showed the closest homologue with functional annotation for SEQ ID NO: 1-245 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 1-245 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), polypeptides encoded by SEQ ID NO: 1-245 (i.e. SEQ ID NO: 246-490) were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) polypeptides encoded by SEQ ID NO: 1-245 (i.e. SEQ ID NO: 246-490) were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The GeneAtlas™ software package (Molecular Simulations Inc. (MSI), San Diego, CA) was used to predict the three-dimensional structure models for the polypeptides encoded by SEQ

ID NO 1-216 (i.e. SEQ ID NO: 246-490). Models were generated by (1) PSI-BLAST which is a multiple alignment sequence profile-based searching developed by Altschul et al, (Nucl. Acids. Res. 25, 3389-3408 (1997)), (2) High Throughput Modeling (HTM) (Molecular Simulations Inc. (MSI) San Diego, CA,) which is an automated sequence and structure searching procedure (http://www.msi.com/), and (3) SeqFold™ which is a fold recognition method described by Fischer and Eisenberg (J. Mol. Biol. 209, 779-791 (1998)). This analysis was carried out, in part, by comparing the polypeptides of the invention with the known NMR (nuclear magnetic resonance) and x-ray crystal three-dimensional structures as templates. Table 5 shows, "PDB ID", the Protein DataBase (PDB) identifier given to template structure; "Chain ID", identifier of the subcomponent of the PDB template structure; "Compound Information", information of the PDB template structure and/or its subcomponents; "PDB Function Annotation" gives function of the PDB template as annotated by the PDB files (http:/www.rcsb.org/PDB/); start and end amino acid position of the protein sequence aligned; PSI-BLAST score, the verify score, the SeqFold score, and the Potential(s) of Mean Force (PMF). The verify score is produced by GeneAtlas™ software (MSI), is based on Dr. Eisenberg's Profile-3D threading program developed in Dr. David Eisenberg's laboratory (US patent no. 5,436,850 and Luthy, Bowie, and Eisenberg, Nature, 356:83-85 (1992)) and a publication by R. Sanchez and A. Sali, Proc. Natl. Acad. Sci. USA, 95:13597-12502. The verify score produced by GeneAtlas normalizes the verify score for proteins with different lengths so that a unified cutoff can be used to select good models as follows:

Verify score (normalized) = (raw score – 1/2 high score)/(1/2 high score)

The PFM score, produced by GeneAtlas™ software (MSI), is a composite scoring function that depends in part on the compactness of the model, sequence identity in the alignment used to build the model, pairwise and surface mean force potentials (MFP). As given in Table 5, a verify score between 0 to 1.0, with 1 being the best, represents a good model. Similarly, a PMF score between 0 to 1.0, with 1 being the best, represents a good model. A SeqFold™ score of more than 50 is considered significant. A good model may also be determined by one of skill in the art based all the information in Table 5 taken in totality.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determined from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication "

Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al, as reference, were obtained for the polypeptide sequences. Table 6 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

Table 7 correlates each of SEQ ID NO: 1-245 to a specific chromosomal location.

Table 8 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-245, and their corresponding priority full length nucleotide sequences in the priority application USSN 09/654,935, the contents of which is incorporated herein by reference in its entirety.

TABLE 1

Tissue Origin	Tissue/RNA Source	Library Name	SEQ ID NO:
adult brain	GIBCO	AB3001	8 24 38 42 56 63-64 93-94 113 130 183 195-196 206 210 227 233 236 240
adult brain	GIBCO	ABD003	2-4 15 19-21 29 31-32 34-39 41-43 45 54 56 67 80 82 84 88 94 103-104 107 113 117 130-131 154 159 178 195 199 206 210 220-221 223
adult brain	Clontech	ABR001	2-3 17 33 35 43 56 62 67 84 113 191 220
adult brain	Clontech	ABR006	2-4 34 82 89 101-102 113 127 146 152 158 162 181 191 197-198 200-201 214 221-223 234 241
adult brain	Clontech	ABR008	2-4 9-12 15 17 19 21 24 29 36-41 54 64 70 74-75 77 79-80 82 84 93-94 97-98 101-102 104 107 109 117 121-124 127 131 140 143-144 146 148-149 151-152 155 158 162 164 167 169 178 193 196 200-202 204 206 221 223-225 227 229 233
adult brain	BioChain	ABR012	2-3 54
adult brain	BioChain	ABR013	17 43 209 240
adult brain	Invitrogen	ABR014	23 43 227 232
adult brain	Invitrogen	ABR015	43 54 65 67 89 142 159 232
adult brain	Invitrogen	ABR016	2-3 28 54 56 64 104 159 229
adult brain	Invitrogen	ABT004	2-3 23 30 33 36-38 40 100 145 152 154 177 191 206 220 242
cultured preadipocytes	Stratagene	ADP001	2-3 15 29 36 38 40 43 56 100 104-105 130 142-144 158-159 177 182 206 236 240
adrenal gland	Clontech	ADR002	11-12 19-20 28 37-38 42 50 56 70 76 82 84 102 104-105 127 130 145 148-150 181 183 189 191 209-210 224-225
adult heart	GIBCO	AHR001	2-5 8-9 11-12 19-22 24 29 36 38 40 43 45 47 54 56 62-63 70 72 74 76 79 82 84 86 92 94 101-104 107 113 127 130-131 137-138 140 143-144 148-149 159 166 169 177-178 183 196 206-207 210 214 229-233 236-237
adult kidney	GIBCO	AKD001	2-3 7-9 11-12 15 18 20-21 24 26-27 29 31-33 36-43 52 54 56 61-62 64 80 82 91 95 98 101-104 107 113 117 130-131 143-144 146 154 159 169 178 181 183 191 195-199 204 206 210 214 220 223-225 227 229 233 240 244
adult kidney	Invitrogen	AKT002	6 8-9 11-12 18 33 36-37 40 43 46 56 64 82 84 86-87 91 107 113 130 142 144 148-149 152 159 167 169 183 191 193 206 223 226 228 232 240-241 244
adult lung	GIBCO	ALG001	5 15 20 29 43 47 54 56 88 103 130 173 177 183 191 214 232 240 244
lymph node	Clontech	ALN001	8 29 36 46 104 130 159 183 206 214 240
young liver	GIBCO	ALV001	2-3 11-12 15 19 37-38 40 43 47 56 62 70 94 103 107 112 143-144 162 181 183 191 195 206 214 220 224-225 236-237 243
adult liver	Invitrogen	ALV002	2-3 10-12 15 20 22 26-27 37 50 89 143 148-149 173 181 183 191 193 206 217 220 240 244
adult liver	Clontech	ALV003	21 181 232
adult ovary	Invitrogen	AOV001	2-3 8 10-12 14-15 19-23 26-29 31-32 34 36-43 47 50 56 62-64 67 70 75 78 82 84 86 89 94 101-102 104 107 109 113 118 125 130-131 140 142 144 146 148-150 152 155 158-159 162 166-167 169 173 177-178 182-183 189 193 195 204 206 210 214 223-225 227 232 240-244
adult placenta	Clontech	APL001	43 159 169 206 240
placenta	Invitrogen	APL002	20 26-27 36 38 64 71 100 178 196 220 228 233
adult spleen	GIBCO	ASP001	2-3 8 26-27 29 35 37 42-43 46-47 54 56 62 64 87 94 104 130 143-144 152 159 183 199 206 214 220 227 232 236 244
adult testis	GIBCO	ATS001	5 8 11-12 20 23-24 29 31-32 37-38 41 43 54 56 62 64 86 89 104 107 130-131 137-138 159 178 183 195 210 229 232 236-237
adult bladder	Invitrogen	BLD001	8 54 159 195 206

Tissue Origin	Tissue/RNA Source	Library Name	SEQ ID NO:
bone marrow	Clontech	BMD001	2-5 8-12 19 22 26-27 29 31-32 34 36-38 42-43 46-47 56 63-64 70 80 86-87 89 91 93-94 98 103-104 107 109 113 118 130-131 144 146 152 159 162 167 178 182 193 199 206-207 210 214 220 223 228 232 240 244
bone marrow	Clontech	BMD002	2-3 5 8 11-12 15 21 26-27 29 36 40 42 45-46 50 54 56 91 94 97-98 104-105 107 109 120 124 137-138 140 142 144 159 165 167 169 173 183 189 191 193 196 204-206 226 232-234 236-237 244
bone marrow	Clontech	BMD004	232
bone marrow	Clonetech	BMD007	43 232
adult colon	Invitrogen	CLN001	38 43 45-46 50 84 87 143 193 195 222 244
mixture of 16 tissues-mRNAs*	various vendors	CTL016	20
mixture of 16 tissues-mRNAs*	various vendors	CTL021	46 54 159 232
mixture of 16 tissues-mRNAs*	various vendors	CTL028	159 237
adult cervix	BioChain	CVX001	2-3 8 11-12 15 21 24 31-32 35-36 39-43 46 56 62-65 70 82 87 89 93-94 98 105 107 120 125-126 131 144 148-150 152 159 165 178 182-183 189 191 193 195 223 236 240
endothelial cells	Strategene	EDT001	2-4 8 10-12 15 21-24 28-30 33-34 36-37 40 42-43 45 47 50 56 62 64 67 70 72 80 82 86 94 103-104 107 109 126 130-131 142-144 146 148-149 152 154 158-159 162 169 177-178 182-183 191 193 195-199 206 210 214 223-226 229 233 236 240-242
fetal brain	Clontech	FBR001	43 130 199
fetal brain	Clontech	FBR004	31-32
fetal brain	Clontech	FBR006	2-4 8 10 29 39 41 43 49 70 77 80 82 84 89 94 104-105 118 121-123 142 150-152 154-155 165 178 186 200-201 204 206-207 210
fetal brain	Invitrogen	FBT002	2-3 8 11-12 29 37 43 67 82 89 134 142-143 152 159 177 189 191 193 199 206 210 220 227
fetal heart	Invitrogen	FHR001	41
fetal kidney	Clontech	FKD001	2-3 10-12 17 29 38 40 43 54 69 75 80 127 159 229 231 236 240
fetal kidney	Clontech	FKD002	56
fetal kidney	Invitrogen	FKD007	19 36 43 56 159
fetal lung	Clontech	FLG001	2-3 54 69 109 113
fetal lung	Invitrogen	FLG003	10 21 35 43 50 54 69 80 92 125-126 143 148-149 158-159 199 221 231-232
fetal liver-spleen	Columbia University	FLS001	1-5 7-12 14-15 18-24 26-28 30 36-38 40-43 50 54 56 62 64 70 72 75 82 84 86 89 91 94-95 98 100 102-105 107 109 112-113 121 130-131 137-138 140 142-144 146 151-152 158-159 162 165-166 169 177-178 181 183 189 191 193 195-198 204-206 210 214 216 220 223-228 230-233 236-237 240-241 244
fetal liver-spleen	Columbia University	FLS002	1-4 6 10-12 14-15 17-18 20-22 29-30 33 36 38-40 42 45 56 62-64 70 75 80 82 91-92 94-95 98 103-105 109 112-113 121 126 131 142 144 146 148-149 152 162 165-167 169 181 183 186 189 191 193 195-199 205-207 214 223 227-228 233
fetal liver-spleen	Columbia University	FLS003	94 112 167 181 183 185 223 232
fetal liver	Invitrogen	FLV001	1-3 6-8 15 18 23 36-39 43 62 80 82 143 145 152 177 181 191 195 206 232

Tissue Origin	Tissue/RNA Source	Library Name	SEQ ID NO:
fetal liver	Clontech	FLV004	2-3 22 24 36 82 109 122-123 152 162 181 232
fetal muscle	Invitrogen	FMS001	5 28 43 47 56 72 78-79 100 137-138 144 152 154 159 169 193 207 210 237 241
fetal muscle	Invitrogen	FMS002	5 137-138 241
fetal skin	Invitrogen	FSK001	2-3 8 10 21 35-36 40 43 54 56 62-63 65 69 71 80 84 91 104-105 124 130 132 137-138 142-143 148-151 158-159 166 177-178 182 185 197-198 200-201 206 210 217 230 232 241
fetal skin	Invitrogen	FSK002	2-3 8 11-12 21 24 26-27 29 40 43 50 62 82 88 94 98 104 107 142 148-149 169 185 193 195 216 237
fetal spleen	BioChain	FSP001	183
umbilical cord	BioChain	FUC001	2-3 5 7-8 15 20 26-27 31-32 34 36 38-40 43 45 50 54 56 62 76 82 84 94 103-105 107 121-123 130 143-144 146 148-149 152 154 158-159 178 193 197-198 210 227 232 237 240
fetal brain	GIBCO	HFB001	2-3 8 10-12 15 20-22 24 28-29 31-33 36-38 41 43 54 62 64 67 70 82 88-89 93 98 101-104 107 109 113 117 130-131 140 142 144-145 162 167 178 182-183 189 193 195 197-199 207 210 223 227 229 232
macrophage	Invitrogen	HMP001	8 169
infant brain	Columbia University	IB2002	2-3 9-12 15 20-21 23-24 33-34 38 41-43 49 56 63-64 84 89 100 104-105 107 113 118 146 148-150 152 154-155 158 162 165-166 173 177-178 182 191 193 195 197-201 206 223 227 230-231 237 241
infant brain	Columbia University	IB2003	2-3 11-12 17 100 113 150 158 166 178 191 220-221 223 227
infant brain	Columbia University	IBM002	43 117 173
infant brain	Columbia University	IBS001	23 29 54 94 109 166 220
fibroblast	Strategene	LFB001	2-3 8 11-12 19 29 36-37 43 45 54 56 104-105 113 130 148-149 154 159 169 178 182-183 214 236 240
lung tumor	Invitrogen	LGT002	2-3 5-6 8 11-12 20-22 24 38 40-41 43 46 52 54 56 62 64-65 70 72 80 82 87 89 93 100 104 107 130-131 140 142-145 152 154 159 162 167 177 182-183 195 197-199 206 210 214 223 236 244
lymphocytes	ATCC	LPC001	2-3 11-12 20 22 38 42 50 54 73 80 86 89 94 97 105 127 145 159 162 177 206 213-214 232 234
leukocyte	GIBCO	LUC001	2-4 8 10-12 15 17 19-22 24 26-27 29 35-38 40-43 47 54 56 62 64 70 72 80 82 84 86 89 91 93-94 101-102 104-105 107 109 130-131 143-144 146 154 158-159 162 165 167 169 177-178 182-183 189 191 193 195 200-202 204 206 210 214 217 223 228-229 231-232 236 240-242
leukocyte	Clontech	LUC003	20 42 80 94 105 140 165 191 205 207 214 231
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	42-43 56 64 82 103 107 130 202 206 214 224-225 229 240
mammary gland	Invitrogen	MMG001	2-4 8-9 11-12 15 17 21 26-27 35-36 38-40 43 46 56 61 64-65 71 80 84 87 89 92 94-95 100-102 107 125 131-132 137-138 140 143 145 150 152 154 159 162 166 169 173 177 182-183 191 193 195 197-199 206 210 224-225 227 237 243-244
induced neuron cells	Strategene	NTD001	2-3 29 34 43 45 54 70 89 159 224-225
retinoic acid-induced neuronal cells	Strategene	NTR001	20 124 130 150 152 178 202 217
neuronal cells	Strategene	NTU001	40 43 47 72 131 217 237
pituitary gland	Clontech	PIT004	15 37-38 43 56 130-131 240

Tissue Origin	Tissue/RNA Source	Library Name	SEQ ID NO:
placenta	Clontech	PLA003	2-3
prostate	Clontech	PRT001	5 11-12 43 62 65 83 103 134 152 232 237
rectum	Invitrogen	REC001	2-3 15 18 26-27 43 54 56 73 80 130 145 152 183 199 244
salivary gland	Clontech	SAL001	14 17 29 43 47 70 98 104 132 159 178 196 204 232-233 236-237
salivary gland	Clontech	SALs03	37 137-138 244
skin fibroblast	ATCC	SFB001	43 47
skin fibroblast	ATCC	SFB002	54
skin fibroblast	ATCC	SFB003	100
small intestine	Clontech	SIN001	21 34 46 73-74 86 103 107 130 137-138 144 169 183 193 227-228 237 242-244
skeletal muscle	Clontech	SKM001	5 20 45 79 86 137-138 152 206
skeletal muscle	Clontech	SKM002	137-138
skeletal muscle	Clonetech	SKMS03	137-138
skeletal muscle	NULL	SKMS04	137-138
spinal cord	Clontech	SPC001	29 40 43 54 69 75 88-89 91 152 159 162 178 191 195 206 210 223 229 232
adult spleen	Clontech	SPLc01	6 46 50 70 130 140 152 216 240
stomach	Clontech	STO001	18 21 63 67 71 107 159 210 220 229 241 244
thalamus	Clontech	THA002	9 21 42 45 89 100 117 162 183 220 226-227 242
thymus	Clonetech	THM001	2-3 8 11-12 15 21 23-24 29 38-40 43 46 67 80 82 105 131 151 159 162 191 214 244
thymus	Clontech	THMc02	2-4 10-12 22 26-27 31-32 38 43 47 50 54 80 92 94 101-102 127 134 144 146 152 154-155 158-159 162 167 178 182-183 191 193 195-196 200-201 205 210 214 216 218 233 237 240
thyroid gland	Clontech	THR001	2-3 5 8 10-12 17-18 20-21 23-24 29 38 42-43 45 49 54 56 61-62 64 67 70 75-76 78 84 91-92 94 103-105 107 109 122-123 130 134 143 148-149 155 162 167 169 178 182-183 186 191 193 195-198 200-201 214 229 232-233 237 240 244
trachea	Clontech	TRC001	2-3 15 19 36-37 40 47 54 65 72 89 95 107 204-205 210 232 237 244
uterus	Clontech	UTR001	8 31-32 54 56 178 183 206 232 236 243

*The 16 tissue-mRNAs and their vendor source, are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) normal adult kidney mRNA (Invitrogen), 3) normal adult liver mRNA (Invitrogen), 4) normal fetal brain mRNA (Invitrogen), 5) normal fetal kidney mRNA (Invitrogen), 6) normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) human bone marrow mRNA (Clontech), 10) human leukemia lymphablastic mRNA (Clontech), 11) human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

TABLE 2

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
246	AF145657	Drosophila melanogaster	BcDNA.GH10120	728	38
247	X58141	Homo sapiens	mRNA for erythrocyte adducin alpha subunit.	3826	99
248	L29296	Homo sapiens	(clone: SS20B/E6.0) alpha-adducin gene, exons 14, 15, 16.	3387	99
249	AAB63963	Homo sapiens	26-MAR-2001 26-MAY-2000 Human prostate cancer associated antigen protein sequence SEQ ID NO:1325.	1095	97
250	M29458	Homo sapiens	carbonic anhydrase III gene, exon 7.	1441	100
251	AJ006529	Gallus gallus	putative phosphatase	867	60
252	Y08302	Homo sapiens	mRNA for MAP kinase phosphatase 4.	1996	100
253	X53280	Homo sapiens	BTF3a mRNA.	1048	100
254	AB013790	Ateles belzebuth	immunoglobulin alpha heavy chain	74	43
255	AK027387	Homo sapiens	FLJ14481 fis, clone MAMMA1002351, highly similar to Mus musculus dynactin subunit p25 (p25) mRNA.	964	100
256	AK001686	Homo sapiens	FLJ10824 fis, clone NT2RP4001086.	3013	93
257	AK001686	Homo sapiens	FLJ10824 fis, clone NT2RP4001086.	4089	98
258	AK026076	Homo sapiens	FLJ22423 fis, clone HRC08678.	689	100
259	AY037207	Arabidopsis thaliana	AT3g22240/MMP21_1	66	31
260	AAW58394	Homo sapiens	14-SEP-1998 09-OCT-1997 Human spermidine/spermine N1-acetyltransferase.	797	92
261	AF220051	Homo sapiens	hematopoietic stem/progenitor cells protein MDS031 mRNA, complete cds.	844	98
262	AB017563	Homo sapiens	gene, exon 10 and complete cds.	2283	100
263	J03910	Homo sapiens	(clone 14VS) metallothionein-IG (MT1G) gene, complete cds.	367	98
264	X56351	Homo sapiens	ALAS1 (ALASH) mRNA for delta-aminolevulinate synthase (housekeeping) (EC 2.3.1.37).	3333	100
266	U79241	Homo sapiens	clone 23759 mRNA, partial cds.	2304	100
267	AF068291	Homo sapiens	mRNA, partial cds.	699	99
268	BC007235	Homo sapiens	clone MGC:15430, mRNA, complete cds.	398	100
269	X69151	Homo sapiens	mRNA for subunit C of vacuolar proton-ATPase V1 domain.	1958	100
270	AF271784	Homo sapiens	mRNA, complete cds.	1017	92
271	AB025220	Homo sapiens	mRNA for p40phox, complete cds.	1737	100
272	AB025220	Homo sapiens	mRNA for p40phox, complete cds.	1644	96
273	BC001426	Homo sapiens	Similar to ubiquinol-cytochrome c reductase hinge protein, clone MGC:1361, mRNA, complete cds.	346	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
274	AL050051	Homo sapiens	cDNA DKFZp566D193 (from clone DKFZp566D193); partial cds.	481	98
275	BC002517	Homo sapiens	Pirin, clone MGC:2083, mRNA, complete cds.	1543	100
276	X69962	Homo sapiens	FMR-1 mRNA.	2384	100
277	L29074	Homo sapiens	X mental retardation syndrome protein (FMR1) gene, alternative splice products, complete cds; and pseudogene, complete sequence.	2144	92
278	AK001711	Homo sapiens	FLJ10849 fis, clone NT2RP4001414, highly similar to SEPTIN 2 HOMOLOG.	2179	99
279	AK027641	Homo sapiens	FLJ14735 fis, clone NT2RP3002054.	651	99
280	BC009256	Homo sapiens	clone MGC:14860, mRNA, complete cds.	1065	94
281	AL110239	Homo sapiens	cDNA DKFZp566E144 (from clone DKFZp566E144); complete cds.	1234	99
282	BC008714	Homo sapiens	prostatic binding protein, clone MGC:8531, mRNA, complete cds.	1017	100
283	BC004374	Homo sapiens	ARP1 (actin-related protein 1, yeast) homolog B (centractin beta), clone MGC:10568, mRNA, complete cds.	1949	100
284	AF201334	Homo sapiens	mRNA, complete cds.	2395	100
285	BC008743	Homo sapiens	zyxin, clone MGC:3071, mRNA, complete cds.	3145	100
286	BC005957	Homo sapiens	solute carrier family 25 (mitochondrial carrier; peroxisomal membrane protein, 34kD), member 17, clone MGC:14604, mRNA, complete cds.	1557	100
287	AF273053	Homo sapiens	tumor antigen se89-1 mRNA, complete cds.	3570	82
288	AB028893	Homo sapiens	U32, U33, U34, U35, RPS11, U35 genes for ribosomal protein L13a and S11, U32, U33, U34, U35, and U35 snoRNA, complete cds and sequence.	595	100
289	AC003973	Homo sapiens	from chromosome 19, BAC 33152, complete sequence.	5273	81
290	AF253978	Homo sapiens	mRNA, partial cds.	487	85
291	AF018265	synthetic construct	immunoglobulin lambda light chain	278	79
292	BC005134	Homo sapiens	Similar to ribosomal protein L14, clone MGC:11208, mRNA, complete cds.	1102	99
293	AK000869	Homo sapiens	FLJ10007 fis, clone HEMBA1000193.	2635	100
294	AAB7322 9	Homo sapiens	11-MAY-2001 11-AUG-2000 Human phosphatase MTMR7_h.	2127	98
295	BC003618	Homo sapiens	Similar to putative nuclear protein, clone MGC:1819, mRNA, complete cds.	3042	100
296	AAB5434 6	Homo sapiens	09-MAR-2001 08-MAR-2000 Human pancreatic cancer antigen protein sequence SEQ ID NO:798.	4092	99
297	AK000330	Homo sapiens	FLJ20323 fis, clone HEP09648.	2229	100
298	AF176701	Homo sapiens	protein FBL9 mRNA, partial cds.	1072	100
299	X54977	Bos taurus	17,000 dalton myosin light chain	789	100
300	AL096746	Homo sapiens	cDNA DKFZp586E1322 (from clone DKFZp586E1322); partial cds.	1186	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
301	BC000502	Homo sapiens	ribosomal protein L17, clone MGC:8457, mRNA, complete cds.	970	100
302	AC004079	Homo sapiens	clone RP1-167F23 from 7p15, complete sequence.	1965	100
303	X92485	Plasmodium vivax	pval	149	55
304	AK006347	Mus musculus	putative	429	86
305	AL137544	Homo sapiens	cDNA DKFZp434A1520 (from clone DKFZp434A1520); partial cds.	974	98
306	AC006276	Homo sapiens	19, cosmid R28379, complete sequence.	900	99
307	AK024297	Homo sapiens	FLJ14235 fis, clone NT2RP4000167.	2325	100
308	AK005941	Mus musculus	putative	460	88
309	AF265440	Homo sapiens	mRNA, complete cds.	1413	100
311	AB027251	Homo sapiens	for zinc finger protein (ZFD25), complete cds.	4369	100
312	AK008240	Mus musculus	putative	455	100
313	AAB75337	Homo sapiens	03-APR-2001 01-JUN-2000 Human secreted protein sequence encoded by gene 47 SEQ ID NO:156.	138	60
314	AF321191	Homo sapiens	(PRX) mRNA, complete cds, alternatively spliced.	7312	99
315	AF225417	Homo sapiens	kDa protein mRNA, complete cds.	3701	99
316	AK000265	Homo sapiens	FLJ20258 fis, clone COLF7250.	2797	97
317	D90070	Homo sapiens	ATL-derived PMA-responsive (APR) peptide mRNA.	278	100
318	U79725	Homo sapiens	A33 antigen precursor mRNA, complete cds.	1678	100
319	M83679	Rattus norvegicus	RAB15	1077	97
320	AK024715	Homo sapiens	FLJ21062 fis, clone CAS01044.	927	98
321	AK000075	Homo sapiens	FLJ20068 fis, clone COL01755.	1729	99
322	AC007954	Homo sapiens	14 clone RP11-493G17 and CTD-2516D11 map 14q24.3, complete sequence.	4243	100
323	Z33905	Homo sapiens	gene for 43kD acetylcholine receptor-associated protein (Rapsyn).	2150	99
324	AF030027	Equine herpesvirus 4	71	118	22
325	AJ291606	Xenopus laevis	gamma tubulin ring protein	2024	55
326	AAB64610	Homo sapiens	22-MAR-2001 01-JUN-2000 Human secreted protein BLAST search protein SEQ ID NO: 120.	197	72
327	AAB53677	Homo sapiens	09-MAR-2001 08-MAR-2000 Human colon cancer antigen protein sequence SEQ ID NO:1217.	694	99
328	AF159055	Homo sapiens	zipper-like protein (LZLP) mRNA, complete cds.	116	79
329	AL160111	Homo sapiens	1 of a novel human mRNA from chromosome 22.	2126	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
330	AF159055	Homo sapiens	zipper-like protein (LZLP) mRNA, complete cds.	130	80
331	AK026264	Homo sapiens	FLJ22611 fis, clone HSI04961.	685	96
332	X57809	Homo sapiens	rearranged immunoglobulin lambda light chain mRNA.	1223	100
333	AAB87440	Homo sapiens	22-MAY-2001 31-AUG-2000 Human gene 32 encoded secreted protein fragment, SEQ ID NO:181.	513	75
334	AK012475	Mus musculus	putative	2259	84
335	AF090930	Homo sapiens	HQ0478 PRO0478 mRNA, complete cds.	146	72
336	AL080196	Homo sapiens	cDNA DKFZp434C212 (from clone DKFZp434C212).	2292	94
337	AK019766	Mus musculus	putative	1288	71
338	X69398	Homo sapiens	mRNA for OA3 antigenic surface determinant.	1632	100
339	AK019305	Mus musculus	putative	506	96
340	AL078630	Mus musculus	573K1.15 (mm17M1-6 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor LIKE) protein))	1023	81
341	AF118078	Homo sapiens	PRO1848	574	100
342	AK005566	Mus musculus	putative	1218	94
343	U71363	Homo sapiens	zinc finger protein zfp6 (ZF6) mRNA, partial cds.	1367	70
344	AK015315	Mus musculus	putative	556	76
345	AF218451	Homo sapiens	substrate p130Cas mRNA, complete cds.	4579	99
346	AF151046	Homo sapiens	HSPC212	1345	87
347	AF151046	Homo sapiens	HSPC212	817	74
348	Z14244	Homo sapiens	coxVIIb mRNA for cytochrome c oxidase subunit VIIb.	426	100
349	BC001037	Homo sapiens	ribosomal protein L35a, clone MGC:1639, mRNA, complete cds.	581	100
351	AAB45018	Homo sapiens	12-FEB-2001 09-MAR-2000 Human secreted protein encoded by gene 41 homologue.	142	57
352	AAY94885	Homo sapiens	12-JUN-2000 22-JUL-1999 Human protein clone HP10550.	540	99
353	AF161557	Homo sapiens	HSPC072	472	100
354	AAG01438	Homo sapiens	06-OCT-2000 21-FEB-2000 Human secreted protein, SEQ ID NO: 5519.	353	92
355	AF161507	Homo sapiens	HSPC158	1197	99
356	AL122111	Homo sapiens	cDNA DKFZp434A1721 (from clone DKFZp434A1721).	2868	99
357	AF349540	Homo sapiens	XIII secreted phospholipase A2 mRNA, complete cds.	1073	100
358	AF274714	Homo sapiens	protein-related protein (ORP1) mRNA, complete cds.	2363	100
359	AAG0379	Homo	06-OCT-2000 21-FEB-2000 Human secreted	222	67

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
	3	sapiens	protein, SEQ ID NO: 7874.		
360	BC000705	Homo sapiens	clone MGC:861, mRNA, complete cds.	908	100
361	AAG0378 9	Homo sapiens	06-OCT-2000 21-FEB-2000 Human secreted protein, SEQ ID NO: 7870.	188	60
362	AAB6281 0	Homo sapiens	02-MAY-2001 06-JUL-2000 Human nervous system associated protein NSPRT3 amino acid sequence.	501	96
363	AF161370	Homo sapiens	mRNA, partial cds.	654	91
364	AK011592	Mus musculus	putative	1245	66
365	AK002154	Homo sapiens	FLJ11292 fis, clone PLACE1009665.	230	64
366	AF159297	Zea mays	extensin-like protein	349	28
367	AF125096	Homo sapiens	HSPC042 protein	137	96
368	AF125096	Homo sapiens	HSPC042 protein	243	98
369	AK001745	Homo sapiens	FLJ10883 fis, clone NT2RP4001946, weakly similar to PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE (EC 2.1.1.77).	1880	99
370	AF151783	Homo sapiens	(MEG3) mRNA, complete cds.	3651	99
371	X16707	Homo sapiens	fra-1 mRNA.	1443	100
372	AF176555	Homo sapiens	anchoring protein 220 mRNA, complete cds.	9783	99
373	X78121	Homo sapiens	mRNA.	3404	100
374	U82670	Homo sapiens	Xq28 psHMG17 pseudogene, complete sequence; and melanoma antigen family A1 (MAGEA1) and zinc finger protein 275 (ZNF275) genes, complete cds.	2513	99
375	AK018726	Mus musculus	putative	670	100
376	BC000187	Homo sapiens	cytochrome c oxidase subunit VIc, clone MGC:1520, mRNA, complete cds.	379	100
377	AAY8754 8	Homo sapiens	18-JUL-2000 03-NOV-1997 Human disease-associated calmodulin protein (DACP-1).	729	100
378	AK003198	Mus musculus	putative	562	100
379	AK000496	Homo sapiens	FLJ20489 fis, clone KAT08285.	333	69
380	AF130079	Homo sapiens	PRO2852	308	74
381	AAY9196 1	Homo sapiens	19-JUL-2000 17-SEP-1999 Human cytoskeleton associated protein 16 (CYSKP-16).	1293	96
382	M15202	Rattus norvegicus	troponin T class IIIa beta	1155	94
383	AF026276	Homo sapiens	skeletal troponin T (TNNT3) gene, complete cds.	1205	94
384	AF090694	Homo sapiens	RNA binding protein (NAPOR-2) mRNA, complete cds.	2519	98
385	BC007655	Homo sapiens	protein phosphatase 1, regulatory (inhibitor) subunit 2, clone MGC:1327, mRNA, complete cds.	1051	100
386	AF161533	Homo sapiens	HSPC048	573	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
387	BC002801	Homo sapiens	p47, clone MGC:3347, mRNA, complete cds.	1812	96
388	AK027878	Homo sapiens	FLJ14972 fis, clone THYRO1000715.	2669	98
389	AF161418	Homo sapiens	HSPC300	378	100
390	AK010720	Mus musculus	putative	105	28
391	X66358	Homo sapiens	mRNA KKIALRE for serine/threonine protein kinase.	1929	99
392	AF290612	Homo sapiens	Q0310 liver nuclear protein mRNA, complete cds.	2246	98
393	U69263	Homo sapiens	precursor, mRNA, complete cds.	4516	99
394	U69263	Homo sapiens	precursor, mRNA, complete cds.	4021	99
395	AK000838	Homo sapiens	FLJ20831 fis, clone ADKA03080.	761	100
396	AK006393	Mus musculus	putative	819	90
397	AF312033	Mus musculus	ASR2A	4584	97
398	BC001904	Homo sapiens	Similar to phosphoglycerate mutase 2 (muscle), clone MGC:2269, mRNA, complete cds.	270	100
399	Y14391	Homo sapiens	for putative GTP-binding protein.	2042	99
400	AF242528	Homo sapiens	finger protein 291 (ZNF291) mRNA, complete cds.	294	100
401	AF116695	Homo sapiens	PRO2221	173	46
402	AAR32020	Homo sapiens	11-JUL-1993 14-AUG-1992 Sequence of a eukaryotic transcription factor (TF).	734	66
403	AB049127	Homo sapiens	mRNA for MAP/microtubule affinity-regulating kinase like 1, complete cds.	2227	73
404	K03250	Rattus norvegicus	ribosomal protein S11	824	100
405	AF144233	Homo sapiens	binding peptide mRNA, partial cds.	328	96
406	AC007055	Homo sapiens	14 clone BAC 201F1 map 14q24.3, complete sequence.	519	100
407	AK001752	Homo sapiens	FLJ10890 fis, clone NT2RP4002071.	5019	99
408	AF090931	Homo sapiens	HQ0483$ PRO0483 mRNA, complete cds.	133	58
409	A28080	Mycobacterium avium subsp. paratuberculosis	34 kDa protein	75	36
410	AL136704	Homo sapiens	cDNA DKFZp566A1524 (from clone DKFZp566A1524); complete cds.	1662	99
411	AL137347	Homo sapiens	cDNA DKFZp761M1511 (from clone DKFZp761M1511); partial cds.	473	100
412	AK027527	Homo sapiens	FLJ14621 fis, clone NT2RP2000079.	1012	100
413	AAG01083	Homo sapiens	06-OCT-2000 21-FEB-2000 Human secreted protein, SEQ ID NO: 5164.	274	96
414	BC009405	Homo sapiens	adenylate kinase 2, clone MGC:15301, mRNA, complete cds.	1094	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
415	U34994	Homo sapiens	dependent protein kinase catalytic subunit (PRKDC) mRNA, complete cds; alternatively spliced.	21178	100
416	U47077	Homo sapiens	protein kinase catalytic subunit (DNA-PKcs) mRNA, complete cds.	21319	99
417	U22229	Felis catus	ribosomal protein L41	128	100
418	AF361481	Homo sapiens	GTP-binding protein 1 (GTPBP3) gene, complete cds; nuclear gene for mitochondrial product.	1402	94
419	BC000606	Homo sapiens	Similar to ribosomal protein L14, clone MGC:1644, mRNA, complete cds.	1094	100
421	AAY7334 5	Homo sapiens	24-FEB-2000 04-MAY-1999 HTRM clone 438283 protein sequence.	2171	73
422	AK000632	Homo sapiens	FLJ20625 fis, clone KAT04008.	816	100
423	AC004668	Homo sapiens	clone CTA-276O3 from 7q22-q31.1, complete sequence.	1976	99
424	AK000496	Homo sapiens	FLJ20489 fis, clone KAT08285.	238	73
425	AAY0278 5	Homo sapiens	11-JUN-1999 07-JUL-1998 Human secreted protein encoded by gene 51 clone HUKEX85.	82	43
426	AF118092	Homo sapiens	PRO2061	1440	96
427	AK000382	Homo sapiens	FLJ20375 fis, clone HUV00942.	1330	99
428	Y15286	Homo sapiens	for vacuolar proton-ATPase subunit M9.2.	459	100
429	AK014098	Mus musculus	putative	524	68
430	AF286095	Homo sapiens	receptor (IL22R) mRNA, complete cds.	629	86
431	AK023266	Homo sapiens	FLJ13204 fis, clone NT2RP3004507, weakly similar to MOB1 PROTEIN.	758	90
432	AF047354	Homo sapiens	and spleen DNase precursor (LSD) mRNA, complete cds.	1046	99
433	X53682	Homo sapiens	LAG-1 gene.	484	100
434	AC000064	Homo sapiens	BAC clone RG083M05 from 7q21-7q22, complete sequence.	298	100
435	AL390921	Arabidopsis thaliana	putative protein	72	44
436	AAB8744 0	Homo sapiens	22-MAY-2001 31-AUG-2000 Human gene 32 encoded secreted protein fragment, SEQ ID NO:181.	1572	100
437	AP003001	Mesorhizobi um loti	O-linked GlcNAc transferase	153	30
438	AK000642	Homo sapiens	FLJ20635 fis, clone KAT03466.	1854	99
439	Z48810	Homo sapiens	mRNA for TX protease precursor.	306	92
441	AC003002	Homo sapiens	DNA from overlapping chromosome 19-specific cosmids R29515 and R28253, genomic sequence, complete sequence.	436	98
442	AF109377	Mus musculus	IdlBp	3979	82
443	AF109377	Mus musculus	IdlBp	2711	81
444	AAG0204 2	Homo sapiens	06-OCT-2000 21-FEB-2000 Human secreted protein, SEQ ID NO: 6123.	797	100

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
445	M17877	Plasmodium falciparum	interspersed repeat antigen	291	27
446	M17877	Plasmodium falciparum	interspersed repeat antigen	291	27
447	AB025784	Rattus norvegicus	PPAR gamma coactivator	331	46
448	AK000755	Homo sapiens	FLJ20748 fis, clone HEP05772.	831	96
449	AK001714	Homo sapiens	FLJ10852 fis, clone NT2RP4001498, weakly similar to ANKYRIN REPEAT-CONTAINING PROTEIN AKR1.	2586	100
450	AB042646	Homo sapiens	mRNA, complete cds.	1224	100
451	AF125533	Homo sapiens	b5 reductase isoform mRNA, complete cds.	1606	100
452	AAY02591	Homo sapiens	19-JUL-1999 09-OCT-1998 A human progesterone receptor complex p23-like protein.	849	100
453	BC000600	Homo sapiens	Similar to from HeLa cyclin-dependent kinase 2 interacting protein, clone MGC:849, mRNA, complete cds.	1106	100
454	Z46937	Caenorhabditis elegans	similarity with ribosomal protein L21	140	38
455	AF161556	Homo sapiens	HSPC071	941	100
456	AF225971	Homo sapiens	(TUBG2) mRNA, complete cds.	2346	99
458	AF343664	Homo sapiens	receptor translocation associated protein 2c (IRTA2) mRNA, complete cds, alternatively spliced.	736	55
459	AF191545	Homo sapiens	mRNA, complete cds.	4141	99
460	AF118082	Homo sapiens	PRO1902	202	58
461	D00531	Oncorhynchus masou	apopolysialoglycoprotein	512	30
462	Z11898	Homo sapiens	OTF3 mRNA encoding octamer binding protein 3A.	1948	100
464	AL162044	Homo sapiens	cDNA DKFZp761L0812 (from clone DKFZp761L0812); partial cds.	220	41
465	AL137301	Homo sapiens	cDNA DKFZp434N1429 (from clone DKFZp434N1429); partial cds.	543	100
466	AB032593	Homo sapiens	for PXR2b, complete cds.	3201	100
467	AL050075	Homo sapiens	cDNA DKFZp566F0546 (from clone DKFZp566F0546); partial cds.	407	100
468	AK000732	Homo sapiens	FLJ20725 fis, clone HEP13903.	1653	99
469	AB049638	Homo sapiens	mRNA for mitochondrial ribosomal protein L11 (L11mt), complete cds.	941	100
470	AB049638	Homo sapiens	mRNA for mitochondrial ribosomal protein L11 (L11mt), complete cds.	737	99
471	AB014772	Homo sapiens	for MOP-3, complete cds.	1722	99
472	AAY59808	Homo sapiens	18-JAN-2000 03-APR-1998 Human normal ovarian tissue derived protein 85.	778	100
473	AF331500	multiple sclerosis associated retrovirus	recombinant envelope protein	1177	92

SEQ ID NO:	Accession Number	Species	Description	Score	% Identity
		element			
474	AF257330	Homo sapiens	protein mRNA, complete cds.	962	96
475	AK000632	Homo sapiens	FLJ20625 fis, clone KAT04008.	809	99
476	M58511	Homo sapiens	iron-responsive element-binding protein/iron regulatory protein 2 (IRE-BP2/IRP2) mRNA, partial cds.	4968	99
477	AF181989	Homo sapiens	beta subunit variant (HBB) mRNA, complete cds.	588	90
478	AC003002	Homo sapiens	DNA from overlapping chromosome 19-specific cosmids R29515 and R28253, genomic sequence, complete sequence.	752	100
479	BC002924	Homo sapiens	clone IMAGE:3956179, mRNA, partial cds.	1221	99
480	AF109146	Homo sapiens	lectin superfamily 6 (CLECSF6) mRNA, complete cds.	958	99
481	BC005374	Homo sapiens	Similar to RIKEN cDNA 1110001E24 gene, clone MGC:12490, mRNA, complete cds.	995	100
482	X75285	Mus musculus	fibulin-2	5621	81
483	AC007954	Homo sapiens	14 clone RP11-493G17 and CTD-2516D11 map 14q24.3, complete sequence.	1342	100
484	AK016295	Mus musculus	putative	116	27
485	AB028893	Homo sapiens	U32, U33, U34, U35, RPS11, U35 genes for ribosomal protein L13a and S11, U32, U33, U34, U35, and U35 snoRNA, complete cds and sequence.	434	100
486	BC003681	Homo sapiens	clone IMAGE:3453235, mRNA, partial cds.	2829	96
487	AK009235	Mus musculus	putative	1648	92
488	AF294900	Homo sapiens	beta-carotene 15,15'- dioxygenase (BCDO) mRNA, complete cds.	2912	100
489	AAB43979	Homo sapiens	08-FEB-2001 08-MAR-2000 Human cancer associated protein sequence SEQ ID NO:1424.	1051	86
490	AF220025	Homo sapiens	motif protein TRIM5 isoform alpha (TRIM5) mRNA, complete cds; alternatively spliced.	1299	95

TABLE 3

SEQ ID NO:	Accession Number	Description	Results*
247	PF00596	Class II Aldolases and Adducin N-terminal domain proteins.	PF00596C 17.24 9.710e-20 217-243 PF00596B 15.07 4.938e-14 180-202 PF00596D 13.89 4.079e-12 297-315
248	PF00596	Class II Aldolases and Adducin N-terminal domain proteins.	PF00596C 17.24 9.710e-20 217-243 PF00596B 15.07 4.938e-14 180-202 PF00596D 13.89 4.079e-12 297-315
250	BL00162	Eukaryotic-type carbonic anhydrases proteins.	BL00162C 17.78 1.000e-40 88-125 BL00162E 14.93 6.478e-34 189-222 BL00162F 22.68 6.727e-30 226-260 BL00162A 22.92 5.179e-26 16-47 BL00162D 15.06 4.960e-22 126-151 BL00162B 21.43 5.345e-17 51-74
252	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 1.196e-11 288-299
253	PD02749	TRANSCRIPTION PROTEIN FACTOR BTF3 REGULATION NUCL.	PD02749B 12.75 1.000e-40 84-120 PD02749C 13.96 3.739e-34 136-170 PD02749A 9.56 6.000e-15 51-64
256	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 8.419e-09 281-301
257	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 8.419e-09 281-301
260	PF00583	Acetyltransferase (GNAT) family.	PF00583A 12.53 3.571e-12 175-186
262	PD01364	MUCIN GLYCOPROTEIN PRECURSOR MEM.	PD01364B 13.94 1.000e-10 336-352
263	PR00860	VERTEBRATE METALLOTHIONEIN SIGNATURE	PR00860B 7.04 2.929e-20 28-42 PR00860C 9.61 1.474e-14 42-52 PR00860A 5.46 9.229e-12 6-19
264	BL00599	Aminotransferases class-II pyridoxal-phosphate attachment sit.	BL00599B 18.93 8.800e-27 278-307 BL00599D 13.25 8.773e-13 411-424 BL00599C 9.13 5.235e-11 334-344
266	PD01769	REDUCTASE PAPS BIOSYNTHESIS PHOSPHOADENO.	PD01769C 21.60 8.393e-18 416-452
271	PR00497	NEUTROPHIL CYTOSOL FACTOR P40 SIGNATURE	PR00497D 11.91 1.176e-28 192-214 PR00497E 10.43 1.123e-26 241-261 PR00497A 6.92 1.136e-24 56-74 PR00497B 4.99 1.125e-23 74-93 PR00497C 8.89 1.100e-21 131-147 PR00497F 8.66 1.138e-15 297-309
272	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 6.538e-11 177-196
276	PF00013	KH domain proteins family of RNA binding proteins.	PF00013 5.78 2.059e-10 268-280
277	PF00013	KH domain proteins family of RNA binding proteins.	PF00013 5.78 2.059e-10 268-280
280	PF00930	Dipeptidyl peptidase IV (DPP IV) N-terminal region.	PF00930J 8.78 4.231e-09 394-415
282	BL01220	Phosphatidylethanolamine-binding protein family proteins.	BL01220B 16.65 1.000e-40 105-146 BL01220C 14.75 5.846e-34 146-174 BL01220A 22.62 3.400e-31 67-98 BL01220D

SEQ ID NO:	Accession Number	Description	Results*
			18.75 5.364e-31 189-221
283	BL00406	Actins proteins.	BL00406B 5.47 1.000e-40 88-143 BL00406C 6.75 1.000e-40 147-202 BL00406D 12.58 7.000e-40 270-325 BL00406E 8.44 6.087e-39 327-377 BL00406A 9.95 6.087e-29 11-46
284	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227C 25.48 7.792e-26 119-171 BL00227D 18.46 2.286e-20 253-307 BL00227B 19.29 4.720e-13 58-113 BL00227A 24.55 4.649e-12 1-35
285	BL00478	LIM domain proteins.	BL00478B 14.79 3.739e-14 463-478 BL00478B 14.79 3.500e-12 405-420 BL00478B 14.79 6.000e-12 530-545
286	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927B 14.66 6.236e-14 146-168
288	BL00783	Ribosomal protein L13 proteins.	BL00783C 22.43 8.071e-20 87-117 BL00783A 14.55 1.600e-19 8-33 BL00783B 12.76 3.500e-12 74-86
289	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.500e-38 422-461
291	DM00031	IMMUNOGLOBULIN V REGION.	DM00031A 16.80 8.364e-11 20-68
292	PD02808	PROTEIN RIBOSOMAL L14 PROBABLE 60.	PD02808A 12.03 3.739e-38 5-42 PD02808B 19.19 8.500e-36 85-120
294	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 2.756e-12 263-274
295	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 8.093e-09 510-564
297	PR00706	PYROGLUTAMYL PEPTIDASE I (C15) FAMILY SIGNATURE	PR00706B 10.56 6.870e-09 74-87
300	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 4.750e-15 40-58
301	BL00464	Ribosomal protein L22 proteins.	BL00464B 28.48 4.960e-35 106-151 BL00464A 29.41 9.700e-23 17-54
302	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.727e-36 158-201
307	BL01113	C1q domain proteins.	BL01113A 17.99 2.558e-09 712-739
310	BL00226	Intermediate filaments proteins.	BL00226D 19.10 9.571e-40 371-418 BL00226B 23.86 4.600e-38 205-253 BL00226C 13.23 9.500e-26 270-301 BL00226A 12.77 4.000e-16 104-119
311	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 5.135e-34 6-45
312	PD01861	PROTEIN NUCLEAR RIBONUCLEOPROTEIN SMALL MRNA RNA.	PD01861A 14.06 4.393e-11 26-50
315	BL00192	Cytochrome b/b6 heme-ligand proteins.	BL00192A 11.90 3.700e-09 96-136
316	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 6.445e-11 661-676
318	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 4.423e-11 103-137

SEQ ID NO:	Accession Number	Description	Results*
319	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 7.455e-13 9-53
321	BL00378	Hexokinases proteins.	BL00378A 19.01 8.375e-09 279-307
323	BL00405	43 Kd postsynaptic protein.	BL00405C 10.15 1.000e-40 65-115 BL00405D 6.60 1.000e-40 123-166 BL00405G 7.78 1.000e-40 226-263 BL00405H 16.83 1.000e-40 263-302 BL00405I 13.75 1.000e-40 302-339 BL00405J 13.28 1.000e-40 339-373 BL00405K 7.57 1.000e-40 373-413 BL00405B 15.33 6.538e-39 26-58 BL00405F 8.07 1.900e-38 195-226 BL00405E 8.84 1.529e-34 166-192 BL00405A 9.73 1.643e-31 2-26
327	BL00048	Protamine P1 proteins.	BL00048 6.39 8.475e-15 24-51 BL00048 6.39 2.918e-14 26-53 BL00048 6.39 5.279e-14 34-61 BL00048 6.39 5.721e-14 32-59 BL00048 6.39 7.197e-14 11-38 BL00048 6.39 8.082e-14 22-49 BL00048 6.39 2.246e-13 10-37 BL00048 6.39 6.677e-13 33-60 BL00048 6.39 7.092e-13 7-34 BL00048 6.39 7.785e-13 8-35 BL00048 6.39 7.923e-13 23-50 BL00048 6.39 1.926e-12 9-36 BL00048 6.39 1.926e-12 31-58 BL00048 6.39 2.456e-12 20-47 BL00048 6.39 6.294e-12 14-41 BL00048 6.39 7.221e-12 25-52 BL00048 6.39 7.750e-12 12-39 BL00048 6.39 9.868e-12 21-48 BL00048 6.39 1.125e-11 19-46 BL00048 6.39 2.375e-11 13-40 BL00048 6.39 6.875e-11 6-33 BL00048 6.39 8.125e-11 36-63 BL00048 6.39 8.250e-11 18-45 BL00048 6.39 8.250e-11 30-57 BL00048 6.39 1.947e-10 5-32 BL00048 6.39 3.605e-10 4-31 BL00048 6.39 4.908e-10 27-54 BL00048 6.39 5.974e-10 42-69 BL00048 6.39 7.039e-10 15-42 BL00048 6.39 7.750e-10 17-44 BL00048 6.39 7.987e-10 39-66 BL00048 6.39 9.526e-10 1-28 BL00048 6.39 1.225e-09 38-65 BL00048 6.39 3.363e-09 16-43 BL00048 6.39 4.038e-09 3-30 BL00048 6.39 5.950e-09 28-55 BL00048 6.39 6.288e-09 29-56 BL00048 6.39 6.400e-09 40-67 BL00048 6.39 6.738e-09 2-29 BL00048 6.39 7.863e-09 35-62
331	PR00221	CAULIMOVIRUS COAT PROTEIN SIGNATURE	PR00221H 12.82 1.217e-09 27-41
332	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 1.529e-14 187-210 BL00290B 13.17 9.000e-12

SEQ ID NO:	Accession Number	Description	Results*
			247-265
334	BL00415	Synapsins proteins.	BL00415N 4.29 8.420e-10 334-378
336	PR00779	INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN RECEPTOR SIGNATURE	PR00779F 14.51 5.147e-09 512-535
338	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 7.158e-10 107-117
339	BL00224	Clathrin light chain proteins.	BL00224B 16.94 8.200e-09 167-220
340	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237B 13.50 1.000e-11 1-23
343	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 5.154e-15 321-334 PD00066 13.92 2.800e-14 237-250 PD00066 13.92 8.800e-14 265-278 PD00066 13.92 3.000e-13 293-306 PD00066 13.92 9.217e-11 209-222
345	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 4.600e-15 20-36
347	BL00563	Stathmin family proteins.	BL00563D 11.38 4.835e-09 279-315
349	BL01105	Ribosomal protein L35Ae proteins.	BL01105A 17.37 1.000e-40 16-61 BL01105B 12.95 1.000e-40 80-120
350	PD02411	PROTEIN TRANSCRIPTION REGULATION NUCLEAR.	PD02411 21.89 2.929e-15 2227-2261
355	BL00464	Ribosomal protein L22 proteins.	BL00464B 28.48 4.908e-10 128-173 BL00464A 29.41 7.045e-09 69-106
358	BL01013	Oxysterol-binding protein family proteins.	BL01013D 26.81 8.000e-26 358-402 BL01013A 25.14 7.231e-21 45-81 BL01013C 9.97 1.000e-13 132-142 BL01013B 11.33 1.000e-11 110-121
366	PD02557	UREASE ACCESSORY PROTEIN UREF NICKEL.	PD02557C 10.85 6.262e-09 29-44
369	BL01279	Protein-L-isoaspartate(D-aspartate) O-methyltransferase signa.	BL01279A 24.27 7.614e-12 67-115
371	PR00042	FOS TRANSFORMING PROTEIN SIGNATURE	PR00042E 9.69 8.200e-25 154-178 PR00042D 8.97 9.735e-24 133-155 PR00042C 8.29 4.549e-21 115-132 PR00042B 10.70 2.983e-20 98-115 PR00042A 10.04 6.400e-20 39-57
373	PR00893	RAB ESCORT (CHOROIDERAEMIA) PROTEIN SIGNATURE	PR00893H 7.37 2.588e-34 411-439 PR00893J 1.42 1.500e-28 565-586 PR00893D 13.14 1.563e-28 114-138 PR00893C 15.10 2.500e-27 94-115 PR00893K 7.01 1.000e-26 600-620 PR00893I 14.97 2.667e-26 543-563 PR00893A 10.55 1.134e-25 45-64 PR00893F 10.78 3.314e-25 294-313 PR00893E 13.94 1.231e-22 213-230 PR00893G 12.88 5.500e-22 351-368 PR00893B 8.07 6.192e-22 75-93
374	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 9.471e-14 508-525 BL00028 16.07 9.100e-13

SEQ ID NO:	Accession Number	Description	Results*
			424-441 BL00028 16.07 2.957e-12 536-553 BL00028 16.07 4.115e-11 340-357 BL00028 16.07 8.269e-11 452-469 BL00028 16.07 4.300e-10 312-329 BL00028 16.07 7.600e-10 480-497
375	PF01020	Ribosomal L40e family.	PF01020 15.00 1.000e-40 80-129
377	PR00450	RECOVERIN FAMILY SIGNATURE	PR00450C 12.22 7.840e-10 86-108 PR00450C 12.22 7.380e-09 52-74 PR00450C 12.22 7.835e-09 16-38
381	PF00992	Troponin.	PF00992B 26.31 4.000e-30 178-213 PF00992A 16.67 2.636e-29 100-135 PF00992C 16.35 2.800e-15 244-262
382	PF00992	Troponin.	PF00992B 26.31 4.000e-30 157-192 PF00992A 16.67 2.636e-29 79-114 PF00992C 16.35 2.800e-15 223-241
383	PF00992	Troponin.	PF00992B 26.31 4.000e-30 162-197 PF00992A 16.67 2.636e-29 84-119 PF00992C 16.35 2.800e-15 228-246
384	PD02784	PROTEIN NUCLEAR RIBONUCLEOPROTEIN.	PD02784B 26.46 8.307e-10 455-498
385	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 9.686e-09 112-147
388	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 3.323e-14 340-374
391	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 6.553e-13 117-136
393	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 9.571e-16 528-546 PR00453B 14.65 5.000e-13 567-582
394	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 9.571e-16 528-546 PR00453B 14.65 5.000e-13 567-582
399	PR00326	GTP1/OBG GTP-BINDING PROTEIN FAMILY SIGNATURE	PR00326A 8.75 1.514e-09 184-205
402	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 1.692e-10 235-248
403	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 1.529e-16 106-154
404	BL00056	Ribosomal protein S17 proteins.	BL00056A 28.90 3.769e-32 75-115 BL00056B 20.86 6.727e-23 123-147
406	BL00150	Acylphosphatase proteins.	BL00150 25.33 1.000e-40 9-56
410	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245D 10.47 5.224e-09 186-198
413	BL00019	Actinin-type actin-binding domain proteins.	BL00019A 12.56 1.000e-13 38-49
414	BL00113	Adenylate kinase proteins.	BL00113B 20.49 5.667e-32 784-828 BL00113D 24.41 2.565e-27 889-920 BL00113C 12.82 2.286e-16 832-847
415	BL00915	Phosphatidylinositol 3- and 4-kinases proteins.	BL00915B 22.78 9.022e-19 3750-3788 BL00915C 22.43 6.250e-18 3873-3912
416	BL00915	Phosphatidylinositol 3- and 4-kinases	BL00915B 22.78 9.022e-19 3750-

SEQ ID NO:	Accession Number	Description	Results*
		proteins.	3788 BL00915C 22.43 6.250e-18 3904-3943
418	PR00326	GTP1/OBG GTP-BINDING PROTEIN FAMILY SIGNATURE	PR00326A 8.75 2.364e-10 186-207
419	PD02808	PROTEIN RIBOSOMAL L14 PROBABLE 60.	PD02808A 12.03 3.739e-38 5-42 PD02808B 19.19 8.500e-36 85-120
421	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.767e-31 26-65
423	BL00143	Insulinase family, zinc-binding region proteins.	BL00143B 14.41 4.115e-13 102-117
426	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 1.000e-40 206-243 BL00514D 15.35 7.000e-16 251-264 BL00514B 16.42 4.000e-15 150-166 BL00514A 11.68 6.885e-12 40-50
427	PR00536	MELANOCYTE STIMULATING HORMONE RECEPTOR SIGNATURE	PR00536G 6.26 2.688e-09 333-342
432	PR00130	DNASE I SIGNATURE	PR00130E 14.66 5.871e-16 146-176 PR00130D 8.65 2.862e-15 116-146 PR00130H 14.38 1.106e-11 229-250 PR00130F 11.23 1.086e-10 176-206 PR00130G 7.22 2.340e-10 206-229 PR00130A 11.39 7.000e-10 31-61
433	PR00437	SMALL CXC CYTOKINE FAMILY SIGNATURE	PR00437C 14.85 4.696e-09 68-87
445	PF00624	Flocculin repeat proteins.	PF00624J 6.21 9.782e-10 429-484
446	PF00624	Flocculin repeat proteins.	PF00624J 6.21 9.782e-10 429-484
447	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 2.256e-09 222-257
449	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 8.515e-10 120-175
450	BL00027	'Homeobox' domain proteins.	BL00027 26.43 1.818e-21 36-79
451	BL00191	Cytochrome b5 family, heme-binding domain proteins.	BL00191K 17.38 4.951e-27 184-228 BL00191J 11.37 6.447e-17 128-150
454	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 8.457e-09 22-39
456	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227B 19.29 1.000e-40 51-106 BL00227C 25.48 1.000e-40 113-165 BL00227D 18.46 1.000e-40 223-277 BL00227A 24.55 2.607e-31 2-36 BL00227F 21.16 4.316e-30 382-436 BL00227E 24.15 2.667e-23 331-366
457	PR00301	70 KD HEAT SHOCK PROTEIN SIGNATURE	PR00301C 8.62 8.875e-11 235-244
458	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 6.870e-09 47-57 DM00179 13.97 8.435e-09 238-248
459	PR00756	MEMBRANE ALANYL DIPEPTIDASE (M1) FAMILY SIGNATURE	PR00756D 10.58 1.529e-21 367-383 PR00756B 14.06 5.737e-16 253-269 PR00756A 12.90 1.237e-13 205-221 PR00756E 11.91 4.094e-13 386-399 PR00756C 11.60 6.108e-11 331-

SEQ ID NO:	Accession Number	Description	Results*
			342
461	PR00648	GPR3 ORPHAN RECEPTOR SIGNATURE	PR00648B 7.41 8.340e-09 1029-1048
462	BL00027	'Homeobox' domain proteins.	BL00027 26.43 5.500e-27 245-288
466	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 2.862e-09 515-536
469	BL00359	Ribosomal protein L11 proteins.	BL00359A 20.66 5.395e-23 20-56 BL00359B 23.07 4.176e-19 66-107 BL00359C 22.18 2.000e-12 123-157
470	BL00359	Ribosomal protein L11 proteins.	BL00359B 23.07 4.176e-19 40-81 BL00359C 22.18 2.000e-12 97-131
473	PF00429	ENV polyprotein (coat polyprotein).	PF00429 31.08 3.195e-12 299-349
476	BL00450	Aconitase family proteins.	BL00450B 42.34 8.393e-30 281-336 BL00450D 21.14 2.800e-18 560-584 BL00450B 42.34 6.400e-12 341-396 BL00450A 13.76 2.406e-11 246-260 BL00450C 11.95 6.657e-10 507-517
477	BL01033	Globins profile.	BL01033A 16.94 7.923e-18 25-47 BL01033B 13.81 1.000e-15 93-105
480	BL00615	C-type lectin domain proteins.	BL00615A 16.68 5.500e-10 78-96 BL00615B 12.25 7.577e-09 178-192
482	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 5.800e-24 1043-1070 BL01177C 17.39 5.333e-19 997-1016 BL01177B 13.61 7.840e-16 703-719 BL01177D 17.50 1.900e-15 1022-1040
487	BL01032	Protein phosphatase 2C proteins.	BL01032H 11.25 8.200e-09 253-266
489	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 1.563e-15 154-177 BL00290B 13.17 9.000e-12 214-232
490	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 5.886e-10 461-483

*Results include in order: accession number subtype; raw score; p-value; position of signature in amino acid sequence

TABLE 4

SEQ ID NO:	Pfam Model	Description	E-value	Pfam Score
247	Aldolase_II	Class II Aldolase and Adducin N-terminal	7.3e-105	361.8
248	Aldolase_II	Class II Aldolase and Adducin N-terminal	7.3e-105	361.8
249	rrm	RNA recognition motif.	8.8e-06	32.6
250	carb_anhydrase	Eukaryotic-type carbonic anhydrase	7.8e-178	604.2
252	DSPc	Dual specificity phosphatase, catalytic doma	3.6e-69	243.2
253	NAC	NAC domain	4.7e-30	113.3
255	hexapep	Bacterial transferase hexapeptide	6.2e-06	33.1
260	Acetyltransf	Acetyltransferase (GNAT) family	2.8e-19	77.5
262	ig	Immunoglobulin domain	5.2e-20	69.5
263	metalthio	Metallothionein	1.3e-22	88.6
264	aminotran_2	Aminotransferases class-II	2.4e-109	376.7
265	IPP_isomerase	Isopentenyl-diphosphate delta-isomerase	1.6e-128	440.4
266	PAPS_reduct	Phosphoadenosine phosphosulfate reductase	6.2e-14	59.7
271	PX	PX domain	7.4e-31	115.9
272	PX	PX domain	7.4e-31	115.9
276	KH-domain	KH domain	7.2e-13	56.2
277	KH-domain	KH domain	7.2e-13	56.2
278	GTP_CDC	Cell division protein	7.6e-119	408.2
280	abhydrolase_2	Phospholipase/Carboxylesterase	0.013	-41.9
282	PBP	Phosphatidylethanolamine-binding protein	7.8e-88	305.2
283	actin	Actin	1e-174	574.6
284	tubulin	Tubulin/FtsZ family	5e-99	342.4
285	LIM	LIM domain containing proteins	4.6e-36	132.3
286	mito_carr	Mitochondrial carrier proteins	1.4e-41	145.5
288	Ribosomal_L13	Ribosomal protein L13	4.1e-56	199.8
289	zf-C2H2	Zinc finger, C2H2 type	5.4e-268	903.7
291	ig	Immunoglobulin domain	0.053	11.5
292	Ribosomal_L14e	Ribosomal protein L14	3.4e-34	127.0
295	PH	PH domain	3.1e-20	77.3
296	Lysyl_hydro	Lysyl hydrolase	0	2058.2
299	efhand	EF hand	0.075	19.5
300	vwa	von Willebrand factor type A domain	2.8e-35	130.6
301	Ribosomal_L22	Ribosomal protein L22p/L17e	4e-67	236.4
302	homeobox	Homeobox domain	4e-34	126.8
309	IF3	Translation initiation factor IF-3	0.00048	15.1
310	filament	Intermediate filament proteins	9.2e-178	604.0
311	zf-C2H2	Zinc finger, C2H2 type	5.6e-143	488.4
312	Sm	Sm protein	5.6e-26	99.7
314	PDZ	PDZ domain (Also known as DHR or GLGF)	0.037	15.2
316	SH3	SH3 domain	3.6e-12	53.9
318	ig	Immunoglobulin domain	1.5e-12	45.5
319	ras	Ras family	5.1e-94	325.8
321	SAM	SAM domain (Sterile alpha motif)	9.9e-10	45.8
323	TPR	TPR Domain	1.1e-12	55.5
329	rrm	RNA recognition motif.	4.7e-09	43.5
332	ig	Immunoglobulin domain	1e-20	71.8
336	VPS9	Vacuolar sorting protein 9 (VPS9) domain	1.1e-30	115.4
338	ig	Immunoglobulin domain	0.0079	14.2
340	7tm_1	7 transmembrane receptor (rhodopsin family)	2.7e-20	66.6
342	Hydrolase	haloacid dehalogenase-like hydrolase	7.9e-28	105.9
343	zf-C2H2	Zinc finger, C2H2 type	5.1e-35	129.8
345	SH3	SH3 domain	2.2e-14	61.2
349	Ribosomal_L35Ae	Ribosomal protein L35Ae	6e-77	269.0

SEQ ID NO:	Pfam Model	Description	E-value	Pfam Score
350	SET	SET domain	1.1e-56	201.7
358	Oxysterol_BP	Oxysterol-binding protein	3.4e-95	329.7
369	PCMT	Protein-L-isoaspartate(D-aspartate) O-methyl	5e-10	1.8
370	PH	PH domain	9.6e-05	22.0
371	bZIP	bZIP transcription factor	3.2e-07	30.8
373	GDI	GDP dissociation inhibitor	7.4e-25	64.8
374	zf-C2H2	Zinc finger, C2H2 type	7.1e-78	272.1
375	ubiquitin	Ubiquitin family	3.7e-61	193.6
377	efhand	EF hand	1.5e-37	138.2
381	Troponin	Troponin	4.7e-42	153.1
382	Troponin	Troponin	4.7e-42	153.1
383	Troponin	Troponin	4.7e-42	153.1
384	rrm	RNA recognition motif.	7.5e-51	182.4
387	UBX	UBX domain	1.5e-25	98.3
388	G-patch	G-patch domain	4.4e-10	46.9
391	pkinase	Eukaryotic protein kinase domain	1.2e-110	381.1
393	EGF	EGF-like domain	3.6e-82	286.4
394	EGF	EGF-like domain	3.6e-82	286.4
398	PGAM	Phosphoglycerate mutase family	6.1e-07	29.2
402	zf-C2H2	Zinc finger, C2H2 type	4e-24	93.6
403	pkinase	Eukaryotic protein kinase domain	1.1e-101	351.3
404	Ribosomal_S17	Ribosomal protein S17	6e-43	148.6
406	Acylphosphatase	Acylphosphatase	8.5e-64	225.4
407	TPR	TPR Domain	1.2e-14	62.1
414	adenylatekinase	Adenylate kinase	1.9e-119	410.3
415	FAT	FAT domain	9.3e-192	650.4
416	FAT	FAT domain	9.3e-192	650.4
418	MMR_HSR1	GTPase of unknown function	0.00015	-32.8
419	Ribosomal_L14e	Ribosomal protein L14	3.4e-34	127.0
421	zf-C2H2	Zinc finger, C2H2 type	5.2e-99	342.3
423	Peptidase_M16	Insulinase (Peptidase family M16)	4.3e-42	153.3
426	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	2.4e-68	238.3
432	DNase_I	Deoxyribonuclease I (DNase I)	1.2e-171	583.6
433	IL8	Small cytokines (intecrine/chemokine), inter	2.3e-33	115.6
437	TPR	TPR Domain	4.4e-08	40.3
440	PDZ	PDZ domain (Also known as DHR or GLGF)	0.038	15.1
445	zf-C2H2	Zinc finger, C2H2 type	2.7e-22	87.5
446	zf-C2H2	Zinc finger, C2H2 type	4.1e-23	90.2
447	rrm	RNA recognition motif.	0.0029	24.3
449	ank	Ank repeat	4.1e-31	116.8
451	Cyt_reductase	FAD/NAD-binding Cytochrome reductase	7.7e-61	215.5
455	Ribosomal_L18p	Ribosomal L18p/L5e family	0.084	-34.1
456	tubulin	Tubulin/FtsZ family	3.4e-283	954.2
457	laminin_G	Laminin G domain	1.1e-51	185.1
458	ig	Immunoglobulin domain	2.7e-23	80.1
459	Peptidase_M1	Peptidase family M1	6.4e-184	533.4
462	pou	Pou domain - N-terminal to homeobox domain	1.3e-48	175.0
466	TPR	TPR Domain	2.4e-30	114.2
469	Ribosomal_L11	Ribosomal protein L11	7.3e-53	189.0
470	Ribosomal_L11	Ribosomal protein L11	7e-40	145.9
473	ENV_polyprotein	ENV polyprotein (coat polyprotein)	1.5e-37	129.4
476	aconitase	Aconitase family (aconitate hydratase)	2e-189	621.7

SEQ ID NO:	Pfam Model	Description	E-value	Pfam Score
477	globin	Globin	5.5e-44	157.8
480	lectin_c	Lectin C-type domain	1.5e-21	85.0
482	EGF	EGF-like domain	1e-22	88.9
487	PP2C	Protein phosphatase 2C	1.1e-13	51.7
489	ig	Immunoglobulin domain	1.8e-20	71.0
490	7tm_1	7 transmembrane receptor (rhodopsin family)	3.1e-13	44.2

TABLE 5

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
252	1mkp		201	344	3e-40			205.21	PYST1; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE
262	1b2w	L	43	241	8.5e-66			67.25	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRYCTURE, GAMMA- 3 INTERFERON, IMMUNE SYSTEM
262	1b6d	A	43	238	3.4e-65			68.72	IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
262	1bj1	L	43	240	6.8e-67			71.40	FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
262	1bog	A	43	241	6.8e-61			67.70	ANTIBODY (CB 4-1); CHAIN: A, B; PEPTIDE; CHAIN: C;	COMPLEX (ANTIBODY/PEPTIDE) POLYSPECIFICITY, CROSS REACTIVITY, FAB-FRAGMENT, PEPTIDE, 2 HIV-1, COMPLEX (ANTIBODY/PEPTIDE)
262	1bz7	A	43	232	8.5e-60			69.74	ANTIBODY R24 (LIGHT CHAIN); CHAIN: A; ANTIBODY R24 (HEAVY CHAIN); CHAIN: B;	IMMUNE SYSTEM ANTIBODY (FAB FRAGMENT), IMMUNE SYSTEM
262	1ce1	L	43	238	5.1e-65			68.83	CAMPATH-1H:LIGHT CHAIN; CHAIN: L; CAMPATH-1H:HEAVY CHAIN; CHAIN: H; PEPTIDE ANTIGEN; CHAIN: P;	ANTIBODY THERAPEUTIC, ANTIBODY, CD52
262	1dfb	L	43	241	8.5e-66			69.59	IMMUNOGLOBULIN 3D6 FAB 1DFB 3	
262	1fvd	A	43	241	6.8e-66			72.66	IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
262	1gc1	L	43	238	1.2e-62			71.86	ENVELOPE PROTEIN GP120; CHAIN: G; CD4; CHAIN: C; ANTIBODY 17B; CHAIN: L, H;	COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB) COMPLEX (HIV ENVELOPE PROTEIN/CD4/FAB), HIV-1

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										EXTERIOR 2 ENVELOPE GP120, T-CELL SURFACE GLYCOPROTEIN CD4, 3 ANTIGEN-BINDING FRAGMENT OF HUMAN IMMUNOGLOBULIN 17B, 4 GLYCOSYLATED PROTEIN
262	1itb	B	149	429	1.2e-22			67.34	INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
262	1mco	H	29	427	3.4e-68			93.46	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION 1MCO 3	
262	1osp	L	43	241	1.7e-59			69.80	FAB 184.1; CHAIN: L, H; OUTER SURFACE PROTEIN A; CHAIN: O;	COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN) OSPA; COMPLEX (IMMUNOGLOBULIN/LIPOPROTEIN), OUTER SURFACE 2 PROTEIN A COMPLEXED WITH FAB184.1, BORRELIA BURGDORFERI 3 STRAIN B31
262	1wio	A	49	408	9e-17			75.16	T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
262	2fgw	L	43	241	1.2e-67			67.57	IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
262	6fab	L	43	241	5.1e-63			68.83	IMMUNOGLOBULIN ANTIGEN-BINDING FRAGMENT OF THE MURINE ANTI-PHENYLARSONATE 6FAB 3 ANTIBODY 36-71, FAB 36-71 6FAB 4	
263	1mhu		32	62	1.4e-17			67.02	METALLOTHIONEIN CD-7 METALLOTHIONEIN-2 (ALPHA	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									DOMAIN) (/NMR$) 1MHUA 2	
263	4mt2		1	62	1.7e-08			126.36	METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2 3	
264	1ax4	A	190	616	5.1e-10			76.11	TRYPTOPHANASE; CHAIN: A, B, C, D;	TRYPTOPHAN BIOSYNTHESIS TRYPTOPHAN INDOLE-LYASE; TRYPTOPHAN BIOSYNTHESIS, TRYPTOPHAN INDOLE-LYASE, PYRIDOXAL 2 5'-PHOSPHATE, MONOVALENT CATION BINDING SITE
264	1bjw	A	212	590	5.1e-58			85.17	ASPARTATE AMINOTRANSFERASE; CHAIN: A, B;	AMINOTRANSFERASE AMINOTRANSFERASE, PYRIDOXAL ENZYME
264	1bs0	A	203	593	3.4e-72			224.70	8-AMINO-7-OXONANOATE SYNTHASE; CHAIN: A;	TRANSFERASE AONS, 8-AMINO-7-KETOPELARGONATE SYNTHASE; PLP-DEPENDENT ACYL-COA SYNTHASE, BIOTIN BIOSYNTHESIS, 8- 2 AMINO-7-OXONANOATE SYNTHASE, 8-AMINO-7-KETOPELARGONATE 3 SYNTHASE, TRANSFERASE
264	1cs1	A	242	640	3.4e-45			79.69	CYSTATHIONINE GAMMA-SYNTHASE; CHAIN: A, B, C, D;	LYASE CGS; LYASE, LLP-DEPENDENT ENZYMES, METHIONINE BIOSYNTHESIS
264	1d7u	A	213	597	1.7e-46			78.45	2,2-DIALKYLGLYCINE DECARBOXYLASE (PYRUVATE); CHAIN: A;	LYASE DGD; ENZYME COMPLEXES, CATALYTIC MECHANISM, DECARBOXYLATION 2 INHIBITOR, LYASE
264	1qgn	A	215	635	6e-67			88.98	CYSTATHIONINE GAMMA-SYNTHASE; CHAIN: A, B, C, D, E, F, G, H;	LYASE METHIONINE BIOSYNTHESIS, PYRIDOXAL 5'-PHOSPHATE, GAMMA- 2 FAMILY, LYASE
264	1tpl	A	209	612	5.1e-06			86.06	LYASE(CARBON-CARBON) TYROSINE PHENOL-LYASE (E.C.4.1.99.2) 1TPL 3	
264	2gsa	A	170	593	1.4e-72			95.88	GLUTAMATE SEMIALDEHYDE AMINOTRANSFERASE; CHAIN: A, B;	CHLOROPHYLL BIOSYNTHESIS GLUTAMATE SEMIALDEHYDE AMINOMUTASE; CHLOROPHYLL BIOSYNTHESIS, PYRIDOXAL-5'-PHOSPHATE, 2 PYRIDOXAMINE-5'-PHOSPHATE, ASYMMETRIC DIMER

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
266	1sur		226	454	3e-31			66.05	PAPS REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE PHOSPHOADENOSINE PHOSPHOSULFATE REDUCTASE; ASSIMILATORY SULFATE REDUCTION, 3-PHOSPHO-ADENYLYL-SULFATE 2 REDUCTASE, OXIDOREDUCTASE
271	1gri	A	7	231	5.1e-22			57.45	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
272	1gri	A	7	231	5.1e-22			57.45	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
273	1be3	H	22	85	7.5e-26			95.55	CYTOCHROME BC1 COMPLEX; CHAIN: A, B, C, D, E, F, G, H, I, J, K;	ELECTRON TRANSPORT UBIQUINOL CYTOCHROME C OXIDOREDUCTASE, COMPLEX ELECTRON TRANSPORT, CYTOCHROME, MEMBRANE PROTEIN
276	1dt4	A	258	304	1.5e-09	-0.52	0.07		NEURO-ONCOLOGICAL VENTRAL ANTIGEN 1; CHAIN: A;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD, RNA-BINDING MOTIF
276	1dtj	C	258	298	3e-06	-0.27	0.75		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
276	1dtj	D	258	298	3e-06	-0.30	0.93		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
276	1vig		258	296	1.3e-06	-0.20	0.82		VIGILIN; 1VIG 5 CHAIN: NULL; 1VIG 6	RIBONUCLEOPROTEIN RNA-BINDING PROTEIN 1VIG 19
276	2fmr		188	252	3.4e-31	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
276	2fmr		188	252	6e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
276	2fmr		188	252	6e-32			96.30	FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
276	1dt4	A	258	304	1.5e-09	-0.52	0.07		NEURO-ONCOLOGICAL VENTRAL ANTIGEN 1; CHAIN: A;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD, RNA-BINDING MOTIF
276	1dtj	C	258	298	3e-06	-0.27	0.75		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
276	1dtj	D	258	298	3e-06	-0.30	0.93		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
276	1vig		258	296	1.3e-06	-0.20	0.82		VIGILIN; 1VIG 5 CHAIN: NULL; 1VIG 6	RIBONUCLEOPROTEIN RNA-BINDING PROTEIN 1VIG 19
276	2fmr		188	252	6e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
276	2fmr		188	252	6e-32			96.99	FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
276	2fmr		188	252	8.5e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
277	1dt4	A	258	304	1.5e-09	-0.52	0.07		NEURO-ONCOLOGICAL VENTRAL ANTIGEN 1; CHAIN: A;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD, RNA-BINDING MOTIF
277	1dtj	C	258	298	3e-06	-0.27	0.75		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
277	1dtj	D	258	298	3e-06	-0.30	0.93		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
277	1vig		258	296	1.3e-06	-0.20	0.82		VIGILIN; 1VIG 5 CHAIN: NULL; 1VIG 6	RIBONUCLEOPROTEIN RNA-BINDING PROTEIN 1VIG 19
277	2fmr		188	252	3.4e-31	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
277	2fmr		188	252	6e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
277	2fmr		188	252	6e-32			96.30	FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										RNA-BINDING PROTEIN, NMR
277	1dt4	A	258	304	1.5e-09	-0.52	0.07		NEURO-ONCOLOGICAL VENTRAL ANTIGEN 1; CHAIN: A;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD, RNA-BINDING MOTIF
277	1dtj	C	258	298	3e-06	-0.27	0.75		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
277	1dtj	D	258	298	3e-06	-0.30	0.93		RNA-BINDING NEUROONCOLOGICAL VENTRAL ANTIGEN 2; CHAIN: A, B, C, D;	IMMUNE SYSTEM KH DOMAIN, ALPHA-BETA FOLD RNA-BINDING MOTIF
277	1vig		258	296	1.3e-06	-0.20	0.82		VIGILIN; 1VIG 5 CHAIN: NULL; 1VIG 6	RIBONUCLEOPROTEIN RNA-BINDING PROTEIN 1VIG 19
277	2fmr		188	252	6e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
277	2fmr		188	252	6e-32			96.99	FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
277	2fmr		188	252	8.5e-32	0.53	1.00		FMR1 PROTEIN; CHAIN: NULL;	RNA-BINDING PROTEIN KH1; FMR1, FRAGILE X, MODULAR PROTEINS, RNA-BINDING PROTEIN, NMR
278	1zbd	A	35	239	6.8e-56	-0.01	0.01		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
278	3rab	A	37	236	3.4e-56	0.14	-0.07		RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
280	1a88	A	225	450	5.1e-20	-0.05	0.03		CHLOROPEROXIDASE L; CHAIN: A, B, C;	HALOPEROXIDASE BROMOPEROXIDASE L, HALOPEROXIDASE L; HALOPEROXIDASE, OXIDOREDUCTASE
280	1azw	A	225	449	1e-21	0.15	-0.13		PROLINE IMINOPEPTIDASE; CHAIN: A, B;	AMINOPEPTIDASE AMINOPEPTIDASE, PROLINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										IMINOPEPTIDASE, SERINE PROTEASE, 2 XANTHOMONAS CAMPESTRIS
280	1brt		239	451	1.7e-20	0.07	0.22		BROMOPEROXIDASE A2; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE A2, CHLOROPEROXIDASE A2; HALOPEROXIDASE, OXIDOREDUCTASE, PEROXIDASE, ALPHA/BETA 2 HYDROLASE FOLD, MUTANT M99T
280	1cqw	A	233	378	5.1e-21	0.22	-0.18		HALOALKANE DEHALOGENASE; 1-CHLOROHEXANE CHAIN: A;	HYDROLASE A/B HYDROLASE FOLD, DEHALOGENASE 1-S BOND
280	1ehy	A	235	447	1.7e-21	0.07	-0.17		SOLUBLE EPOXIDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HYDROLASE, ALPHA/BETA HYDROLASE FOLD, EPOXIDE DEGRADATION, 2 EPICHLOROHYDRIN
280	1ek1	B	220	394	1.7e-22	0.07	-0.08		EPOXIDE HYDROLASE; CHAIN: A, B;	HYDROLASE HOMODIMER, ALPHA/BETA HYDROLASE FOLD, DISUBSTITUTED UREA 2 INHIBITOR
280	1evq	A	232	438	1.7e-20	0.34	0.24		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
280	1qfm	A	157	453	8.5e-33	-0.05	0.01		PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA- 2 PROPELLER
281	1fxx	A	134	302	6.8e-27	-0.08	0.23		EXONUCLEASE I; CHAIN: A;	HYDROLASE EXODEOXYRIBONUCLEASE I; ALPHA-BETA DOMAIN, SH3-LIKE DOMAIN, DNAQ SUPERFAMILY
282	1a44		48	232	3e-83	1.02	1.00		PHOSPHATIDYLETHANOLAMINE -BINDING PROTEIN; CHAIN: NULL;	LIPID-BINDING PROTEIN PEBP, PBP LIPID-BINDING
282	1a44		48	232	3e-83			317.69	PHOSPHATIDYLETHANOLAMINE -BINDING PROTEIN; CHAIN: NULL;	LIPID-BINDING PROTEIN PEBP, PBP LIPID-BINDING
282	1a44		48	232	6.8e-80	1.02	1.00		PHOSPHATIDYLETHANOLAMINE -BINDING PROTEIN; CHAIN: NULL;	LIPID-BINDING PROTEIN PEBP, PBP LIPID-BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
282	1bch	A	49	232	6e-82	1.05	1.00		PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN; CHAIN: A, B;	LIPID-BINDING LIPID-BINDING, SIGNALLING
282	1beh	A	49	232	6e-82			324.00	PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN; CHAIN: A, B;	LIPID-BINDING LIPID-BINDING, SIGNALLING
282	1beh	A	49	232	8.5e-80	1.05	1.00		PHOSPHATIDYLETHANOLAMINE BINDING PROTEIN; CHAIN: A, B;	LIPID-BINDING LIPID-BINDING, SIGNALLING
283	1dga	A	8	376	0	0.95	1.00		ACTIN; CHAIN: A; GELSOLIN; CHAIN: G;	CONTRACTILE PROTEIN ACTIN, GELSOLIN, CYTOSKELETON ORGANIZATION, ACTIN- 2 ASSOCIATED PROTEIN
283	1esv	A	10	376	0	0.87	1.00		GELSOLIN; CHAIN: S; ALPHA ACTIN; CHAIN: A	CONTRACTILE PROTEIN LATRUNCULIN A, GELSOLIN, ACTIN, DEPOLYMERISATION, 2 SEQUESTRATION
283	1yag	A	8	376	0	0.99	1.00		ACTIN; CHAIN: A; GELSOLIN; CHAIN: G;	CONTRACTILE PROTEIN ACTIN-DEPOLYMERIZING FACTOR (ADF); COMPLEX, ACTIN, GELSOLIN, CONTRACTILE PROTEIN
283	1yag	A	8	376	0			413.68	ACTIN; CHAIN: A; GELSOLIN; CHAIN: G;	CONTRACTILE PROTEIN ACTIN-DEPOLYMERIZING FACTOR (ADF); COMPLEX, ACTIN, GELSOLIN, CONTRACTILE PROTEIN
283	2btf	A	7	376	0	0.91	1.00		ACETYLATION AND ACTIN-BINDING BETA-ACTIN-PROFILIN COMPLEX 2BTF 3	
283	2btf	A	9	376	0			414.62	ACETYLATION AND ACTIN-BINDING BETA-ACTIN-PROFILIN COMPLEX 2BTF 3	
284	1tub	A	1	461	0			285.64	TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
284	1tub	A	1	462	0	0.09	1.00		TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
284	1tub	B	1	459	0	0.11	1.00		TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
284	1tub	B	1	459	0			307.13	TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
285	1a7i		384	437	3e-14	0.43	0.58		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1a7i		384	441	6.8e-10	0.31	0.80		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1a7i		443	500	1.5e-16	0.08	0.58		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1a7i		443	501	1.4e-12	-0.13	0.82		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1a7i		504	566	4.5e-11	-0.40	0.24		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1a7i		504	571	1.2e-09	0.38	0.76		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
285	1b8t	A	375	572	1.4e-23			71.26	CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
285	1b8t	A	379	510	1.4e-23	0.01	-0.17		CRP1; CHAIN: A;	CONTRACTILE LIM DOMAIN, CRP, NMR, MUSCLE DIFFERENTIATION, CONTRACTILE
285	1ctl		376	437	1.7e-12	-0.22	0.10		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
285	1ctl		444	510	3.4e-15	-0.26	0.05		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
285	1ctl		504	571	5.1e-13	0.03	0.22		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
285	1cxx	A	381	437	1.7e-11	-0.17	0.41		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										BINDING PROTEIN
285	1cxx	A	443	496	5.1e-13	0.38	0.53		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
285	1cxx	A	501	568	3.4e-12	0.41	0.87		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
285	1iml		382	440	1.4e-10	-0.25	0.41		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1iml		384	451	4.5e-17	0.21	0.22		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1iml		443	510	1.4e-15	-0.13	0.09		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1iml		443	513	3e-20	0.13	0.12		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1iml		502	569	1.5e-12	0.32	0.93		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1iml		502	571	3.4e-11	0.28	0.99		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
285	1zfo		381	410	1.4e-06	-0.13	0.29		LASP-1; CHAIN: NULL;	METAL-BINDING PROTEIN LIM DOMAIN, ZINC-FINGER, METAL-BINDING PROTEIN
285	1zfo		502	535	0.0012	-0.34	0.15		LASP-1; CHAIN: NULL;	METAL-BINDING PROTEIN LIM DOMAIN, ZINC-FINGER, METAL-BINDING PROTEIN
288	1ffk	G	5	114	9e-49	-0.14	1.00		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	PROTEIN L15P; HMAL15; HL9; 50S RIBOSOMAL PROTEIN L18P; HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E; HL29, L19; 50S RIBOSOMAL PROTEIN L19E; HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E; HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P; HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN
288	1ffk	G	7	135	5.1e-32	0.18	1.00		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P; HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN
289	1a1h	A	1023	1104	1.4e-40	0.06	0.98		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	1051	1132	9e-44	0.09	0.84		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	1611	1715	1.2e-39	0.04	0.46		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	1826	1906	1.7e-30	-0.47	0.45		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	1854	1934	6.8e-31	-0.10	0.05		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	559	639	5.1e-27	0.05	0.17		QGSR ZINC FINGER PEPTIDE;	COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	592	668	1.5e-29	0.15	0.11		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	911	992	6e-45	0.22	0.93		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	939	1020	3e-42	0.04	0.72		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	967	1047	4.5e-42	-0.00	0.78		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1a1h	A	995	1075	9e-42	0.21	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
289	1mey	C	1022	1103	1.4e-39	0.28	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1050	1131	1.7e-41	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1078	1159	1.7e-43	0.35	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1106	1187	3.4e-45	0.16	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(ZINC FINGER/DNA)
289	1mey	C	1134	1215	6.8e-47	-0.08	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN: CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1162	1243	5.1e-48	0.45	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1190	1271	1.7e-48	0.44	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1218	1299	1.4e-49	0.22	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1246	1327	1.4e-49	0.05	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1274	1355	3.4e-50	0.29	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1302	1383	3.4e-49	0.04	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1330	1411	1e-47	0.24	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1358	1439	8.5e-47	0.50	1.00		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1386	1467	1.7e-47	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1414	1495	1.2e-48	0.50	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1414	1496	1.4e-49			103.44	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1442	1523	1.4e-49	0.38	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1470	1551	1e-49	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1498	1579	1.7e-49	0.13	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1526	1607	3.4e-49	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1554	1635	1.7e-49	0.26	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(ZINC FINGER/DNA)
289	1mey	C	1582	1663	1.7e-48	0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1610	1686	1.7e-44	0.28	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1666	1742	8.5e-44	0.52	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1689	1770	5.1e-49	0.41	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1717	1798	1.4e-49	0.03	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1745	1822	3.4e-45	-0.22	0.12		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1825	1906	1e-49	-0.28	0.48		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1853	1934	1e-49	-0.20	0.78		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	1881	1938	1.7e-33	0.35	0.58		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	558	639	3.4e-44	-0.04	0.55		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	586	667	3.4e-46	-0.05	0.82		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	614	695	1.4e-47	0.23	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	642	723	8.5e-49	0.03	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	698	779	1e-49	0.11	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	726	807	6.8e-50	0.19	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	754	835	6.8e-50	0.05	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	782	863	1e-49	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(ZINC FINGER/DNA)
289	1mey	C	810	891	3.4e-49	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	838	935	3.4e-44	0.04	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA).
289	1mey	C	866	963	8.5e-41	0.03	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	910	991	3.4e-42	0.16	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	C	994	1075	1.4e-39	0.59	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1mey	G	908	935	1.5e-10	0.46	0.94		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
289	1tf6	A	1051	1196	1.2e-33	0.18	0.86		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1106	1272	1.7e-36			113.59	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
289	1tf6	A	1163	1308	1.7e-36	-0.10	0.86		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1275	1420	6.8e-37	0.07	0.99		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1387	1532	1.4e-36	0.39	0.90		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1443	1588	3.4e-37	0.20	0.76		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1555	1695	1.2e-33	-0.13	0.64		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	1667	1808	1e-33	-0.29	0.25		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	532	676	1.7e-30	0.04	0.17		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	643	788	3.4e-36	0.06	0.95		TFIIIA; CHAIN: A, D; 5S	COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	699	849	6.8e-38	-0.10	0.94		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	811	951	3.4e-30	0.05	0.87		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1tf6	A	867	1033	6.8e-31	-0.12	0.42		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
289	1ubd	C	1020	1131	1.5e-54	0.04	0.94		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1077	1187	1e-55	0.05	0.94		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1104	1244	3e-53	-0.46	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1160	1271	1.5e-52	0.00	0.72		YY1: CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1198	1299	3.4e-34	0.18	0.58		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1216	1327	6e-52	0.06	0.89		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1226	1327	1.4e-34	0.30	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1245	1356	1.2e-52	0.33	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										REGULATION/DNA)
289	1ubd	C	1272	1383	7.5e-50	0.01	0.78		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1328	1439	3e-50	0.26	0.86		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1384	1496	4.5e-52	0.11	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1469	1579	4.5e-55	0.03	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1506	1607	3.4e-34	0.09	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1524	1635	4.5e-49	-0.29	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1562	1663	1e-32	0.04	0.94		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1608	1714	6e-52	0.12	0.31		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION. INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1618	1714	3.4e-30	-0.01	0.49		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1636	1742	7.5e-51	-0.22	0.83		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1674	1770	6.8e-32	-0.19	0.90		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	1725	1822	1.7e-30	-0.13	0.12		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	540	639	6.8e-29	-0.21	0.06		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	561	667	1.5e-31	0.20	0.41		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	584	695	3e-42	-0.19	0.86		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	589	695	3.4e-32	-0.17	0.86		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
289	1ubd	C	619	724	1.5e-47	0.04	0.51		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	640	752	1.2e-52	-0.15	0.77		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	650	751	1.2e-33	-0.06	0.92		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	668	779	7.5e-51	0.01	0.57		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	725	835	7.5e-53	-0.06	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	734	835	1e-33	0.22	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, B;	INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	790	891	8.5e-33	0.26	0.87		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	808	935	9e-53	0.00	0.95		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	818	935	1.2e-31	-0.14	0.92		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	864	991	9e-53	0.04	0.83		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	874	991	1.5e-27	-0.28	0.66		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(TRANSCRIPTION REGULATION/DNA)
289	1ubd	C	964	1076	3e-53	0.10	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
289	2gli	A	1022	1188	3e-72	-0.09	0.96		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1106	1273	7.5e-71	0.05	0.89		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1190	1329	1.3e-67	0.30	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1254	1382	5.1e-34	0.12	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1302	1469	4.5e-67	0.04	0.86		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1386	1525	4.5e-67	0.19	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1414	1580	6e-71	-0.20	0.92		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1498	1716	1.5e-66	-0.17	0.16		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
289	2gli	A	1534	1662	1.7e-32	0.03	0.63		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1554	1744	4.5e-67	-0.17	0.59		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1590	1713	1.7e-30	-0.13	0.78		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1638	1768	4.5e-65	0.18	0.86		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	1646	1797	8.5e-33	0.00	0.62		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	558	694	3.4e-33	-0.28	0.62		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	587	725	1.5e-53	-0.31	0.19		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	614	781	1.5e-63	-0.20	0.49		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	622	753	1e-33	0.11	0.46		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	642	809	1.5e-68	0.01	0.96		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	670	837	3e-66	-0.19	0.84		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	726	893	6e-68	0.00	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	734	862	1.5e-33	0.06	0.82		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	790	934	1.7e-30	-0.04	0.40		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	838	992	1.5e-69	-0.00	0.69		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	874	993	1.7e-26	-0.10	0.81		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	910	1077	4.5e-70	-0.01	0.96		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
289	2gli	A	938	1105	1.5e-69	0.03	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
291	1cic	B	20	66	1.5e-23	-0.58	0.06		IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
291	1fsk	C	20	66	8.5e-22	-0.56	0.00		MAJOR POLLEN ALLERGEN BET V 1-A; CHAIN: A, D, G, J; IMMUNOGLOBULIN KAPPA LIGHT CHAIN; CHAIN: B, E, H, K; ANTIBODY HEAVY CHAIN FAB;	IMMUNE SYSTEM BET V I-A, BETVI ALLERGEN; BV16 FAB-FRAGMENT, KAPPA MOPC21 CODING SEQUENCE; HEAVY CHAIN OF THE MONOCLONAL ANTIBODY MST2;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: C, F, I, L;	BET V 1, BV16 FAB FRAGMENT, ANTIBODY ALLERGEN COMPLEX
291	1jhl	H	20	66	6.8e-22	-0.72	0.09		COMPLEX(ANTIBODY-ANTIGEN) FV FRAGMENT (IGG1, KAPPA) (LIGHT AND HEAVY VARIABLE DOMAINS 1JHL 3 NON-COVALENTLY ASSOCIATED) OF MONOCLONAL ANTI-HEN EGG 1JHL 4 LYSOZYME ANTIBODY D11.15 COMPLEX WITH PHEASANT EGG 1JHL 5 LYSOZYME 1JHL 6	
292	1vsg	A	123	181	0.00075	0.36	0.09		GLYCOPROTEIN VARIANT SURFACE GLYCOPROTEIN (N-TERMINAL DOMAIN) 1VSG 3	
295	1btk	A	30	118	6e-09	0.21	0.07		BRUTON'S TYROSINE KINASE; CHAIN: A, B;	TRANSFERASE BRUTON'S AGAMMAGLOBULINEMIA TYROSINE KINASE, BTK; TRANSFERASE, PH DOMAIN, BTK MOTIF, ZINC BINDING, X-LINKED 2 AGAMMAGLOBULINEMIA, TYROSINE-PROTEIN KINASE
295	1btn		30	110	1.3e-08	0.20	0.25		BETA-SPECTRIN; 1BTN 4 CHAIN: NULL; 1BTN 5	SIGNAL TRANSDUCTION PROTEIN
295	1fb8	A	22	114	1.5e-18	0.62	0.92		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3- CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
295	1fgy	A	9	115	1.5e-14	0.48	0.77		GRP1; CHAIN: A;	SIGNALING PROTEIN ARF1 GUANINE NUCLEOTIDE EXCHANGE FACTOR AND PH DOMAIN
295	1pls		1	115	1.5e-14	0.69	0.95		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									STRUCTURES) 1PLS 5	
295	1pms		33	114	1.5e-11	0.13	0.01		SOS 1; CHAIN: NULL;	SIGNAL TRANSDUCTION SON OF SEVENLESS; PLECKSTRIN, SON OF SEVENLESS, SIGNAL TRANSDUCTION
295	1qqg	A	33	204	3e-18	0.20	-0.14		INSULIN RECEPTOR SUBSTRATE 1; CHAIN: A, B;	SIGNAL TRANSDUCTION IRS-1; BETA-SANDWHICH, SIGNAL TRANSDUCTION
296	1qgq	A	296	467	4.5e-05	-0.21	0.13		SPORE COAT POLYSACCHARIDE BIOSYNTHESIS PROTEIN CHAIN: A;	TRANSFERASE GLYCOSYLTRANSFERASE
297	1erj	A	106	437	1.7e-59	0.03	0.34		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
297	1erj	A	183	481	5.1e-58	0.24	-0.09		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
297	1erj	A	5	251	1.7e-47	-0.04	0.34		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
297	1erj	A	54	352	6.8e-50	0.21	0.95		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
297	1got	B	170	479	3.4e-56	0.24	-0.14		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
297	1got	B	2	252	3.4e-39	-0.24	0.16		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
297	1got	B	31	297	3.4e-44	0.33	0.98		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
297	1got	B	35	369	3.4e-66			59.96	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
297	1got	B	52	349	8.5e-51	0.12	0.96		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
297	1got	B	98	389	3.4e-66	0.28	0.77		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
298	1a4y	A	32	208	7.5e-12	0.44	0.58		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
298	1a4y	A	49	223	1.4e-10	0.05	0.98		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
298	1a9n	A	112	218	0.00051	0.18	0.40		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										SNRNP,RIBONUCLEOPROTEIN
298	1d0b	A	49	219	5.1e-13	0.22	0.64		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
298	1dce	A	53	174	1.7e-07	0.02	0.05		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
298	1ds9	A	113	216	1.2e-09	-0.06	0.04		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
298	1fo1	A	124	210	1.7e-09	0.13	0.37		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
298	1fo1	B	124	210	1.7e-09	0.06	0.13		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
298	1fqv	A	129	214	5.1e-11	0.28	0.46		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
298	1fqv	A	33	140	1.5e-08	0.04	-0.02		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
298	1fqv	A	39	191	3e-21	0.95	1.00		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
298	1fqv	A	49	207	6.8e-19	0.54	0.96		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
298	1fqv	A	70	199	4.5e-19	0.85	0.92		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
298	1fs2	A	129	214	5.1e-11	-0.35	0.27		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
298	1fs2	A	49	207	6.8e-19	0.64	0.77		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
298	1yrg	A	111	220	1e-08	-0.38	0.23		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
298	2bnh		113	223	1.4e-08	0.31	0.76		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
298	2bnh		53	217	1e-10	0.32	1.00		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
299	1brl	B	4	151	3.4e-44	0.91	1.00		MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
299	1brl	B	4	151	3.4e-44			219.13	MYOSIN; CHAIN: A, B, C, D, E, F,	MUSCLE PROTEIN MDE; MUSCLE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									G, H;	PROTEIN
299	1cdm	A	4	149	1.7e-56	0.60	1.00		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
299	1cdm	A	4	149	1.7e-56			103.28	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
299	1cll		4	149	6.8e-62	0.49	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
299	1cll		4	150	6.8e-62			113.59	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
299	1exr	A	4	150	1.4e-59	0.40	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
299	1tcf		3	151	1.7e-48			89.97	TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
299	1tcf		4	148	1.7e-48	0.38	1.00		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
299	1top		4	148	5.1e-49	0.57	1.00		CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
299	1top		4	151	5.1e-49			83.80	CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
299	1vrk	A	2	149	1.2e-60	0.72	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										PROTEIN/PEPTIDE)
299	1vrk	A	2	151	1.2e-60			115.21	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
300	1aox	A	36	215	1.7e-28	0.37	0.83		INTEGRIN ALPHA 2 BETA; CHAIN: A, B;	INTEGRIN INTEGRIN, CELL ADHESION, GLYCOPROTEIN
300	1atz	A	38	226	1.5e-23			72.47	VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN-BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
300	1atz	A	39	218	1.5e-23	0.88	1.00		VON WILLEBRAND FACTOR; CHAIN: A, B;	COLLAGEN-BINDING COLLAGEN-BINDING, HEMOSTASIS, DINUCLEOTIDE BINDING FOLD
300	1auq		23	227	1.4e-35			62.36	A1 DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
300	1auq		29	227	1.4e-35	0.57	1.00		A1 DOMAIN OF VON WILLEBRAND FACTOR; CHAIN: NULL;	WILLEBRAND WILLEBRAND, BLOOD COAGULATION, PLATELET, GLYCOPROTEIN
300	1ck4	A	39	217	5.1e-29	0.58	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
300	1fns	A	36	227	5.1e-34	0.70	0.98		IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: L; IMMUNOGLOBULIN NMC-4 IGG1; CHAIN: H; VON WILLEBRAND FACTOR; CHAIN: A;	IMMUNE SYSTEM VON WILLEBRAND FACTOR, GLYCOPROTEIN IBA (A:ALPHA) BINDING, 2 COMPLEX (WILLEBRAND/IMMUNOGLOBULIN), BLOOD COAGULATION TYPE 3 2B VON WILLEBRAND DISEASE
300	1ido		39	224	5.1e-31			59.31	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
300	1ido		41	217	5.1e-31	0.61	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
300	1lfa	A	38	226	8.5e-23	0.53	0.99		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										ILFA 8
300	1lfa	A	38	227	8.5e-23			53.04	CD11A; ILFA 5 CHAIN: A, B; ILFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; ILFA 8
300	1qc5	A	37	217	1.4e-28	0.41	0.94		ALPHA1 BETA1 INTEGRIN; CHAIN: A; ALPHA1 BETA1 INTEGRIN; CHAIN: B;	CELL ADHESION INTEGRIN, CELL ADHESION
301	1bxe	A	13	153	1.7e-33	-0.14	0.71		RIBOSOMAL PROTEIN L22; CHAIN: A;	RNA BINDING PROTEIN RIBOSOMAL PROTEIN, PROTEIN SYNTHESIS, RNA BINDING, 2 ANTIBIOTICS RESISTANCE, RNA BINDING PROTEIN
301	1ffk	O	2	152	1.7e-44	0.02	1.00		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E: CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O; RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN
302	1ahd	P	143	208	1e-33	-0.16	0.98		DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1AHD 3 REPLACED BY SER (C39S) COMPLEX WITH DNA (NMR, 1AHD 4 16 STRUCTURES) 1AHD 5	
302	1ahd	P	143	209	1e-33			72.79	DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1AHD 3 REPLACED BY SER (C39S) COMPLEX WITH DNA (NMR, 1AHD 4 16 STRUCTURES) 1AHD 5	
302	1b72	A	137	203	1.5e-30			69.28	HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
302	1b72	A	143	203	1.5e-30	-0.07	0.99		HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
302	1b72	A	147	204	1.7e-27	-0.29	1.00		HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
302	1b8i	A	143	202	4.5e-30			61.07	ULTRABITHORAX HOMEOTIC PROTEIN IV; CHAIN: A; HOMEOBOX PROTEIN EXTRADENTICLE; CHAIN: B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA ULTRABITHORAX; PBX PROTEIN; DNA BINDING, HOMEODOMAIN, HOMEOTIC PROTEINS, DEVELOPMENT, 2 SPECIFICITY
302	1b8i	A	144	201	4.5e-30	0.09	0.83		ULTRABITHORAX HOMEOTIC PROTEIN IV; CHAIN: A; HOMEOBOX PROTEIN EXTRADENTICLE; CHAIN: B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA ULTRABITHORAX; PBX PROTEIN; DNA BINDING, HOMEODOMAIN, HOMEOTIC PROTEINS, DEVELOPMENT, 2 SPECIFICITY
302	1ftz		142	210	1.2e-28			71.20	DNA-BINDING FUSHI TARAZU	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PROTEIN (HOMEODOMAIN) (NMR, 20 STRUCTURES) 1FTZ 3	
302	1ftz		144	208	1.2e-28	-0.30	0.59		DNA-BINDING FUSHI TARAZU PROTEIN (HOMEODOMAIN) (NMR, 20 STRUCTURES) 1FTZ 3	
302	1san		148	209	3.4e-31			69.53	DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1SAN 3 REPLACED BY SER AND RESIDUES 1-6 DELETED (C39S,DEL 1-6) 1SAN 4 (NMR, 20 STRUCTURES) 1SAN 5	
302	1san		149	208	3.4e-31	0.00	1.00		DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1SAN 3 REPLACED BY SER AND RESIDUES 1-6 DELETED (C39S,DEL 1-6) 1SAN 4 (NMR, 20 STRUCTURES) 1SAN 5	
302	9ant	A	147	202	5.1e-31	0.38	1.00		ANTENNAPEDIA PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D, E, F;	COMPLEX (DNA-BINDING PROTEIN/DNA) HD; HOMEODOMAIN, COMPLEX (DNA-BINDING PROTEIN/DNA)
302	9ant	A	147	202	5.1e-31			68.47	ANTENNAPEDIA PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D, E, F;	COMPLEX (DNA-BINDING PROTEIN/DNA) HD; HOMEODOMAIN, COMPLEX (DNA-BINDING PROTEIN/DNA)
307	1ddv	A	4	96	0.0003	0.48	0.46		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A; METABOTROPIC GLUTAMATE RECEPTOR MGLUR5; CHAIN: B;	SIGNALING PROTEIN PROTEIN-LIGAND COMPLEX, POLYPROLINE RECOGNITION, BETA TURN
307	1ddw	A	4	96	0.00015	0.62	0.69		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A;	SIGNALING PROTEIN PLECKSTRIN HOMOLOGY DOMAIN FOLD
307	1rrp	B	7	101	1.5e-25	0.57	0.96		RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
309	2ife	A	159	237	1.2e-16	0.62	0.89		TRANSLATION INITIATION FACTOR IF3; CHAIN: A;	GENE REGULATION INITIATION FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
310	1f5n	A	107	167	0.0049	-0.27	0.03		INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1; CHAIN: A;	SIGNALING PROTEIN GBP, GTP HYDROLYSIS, GDP, GMP, INTERFERON INDUCED, DYNAMIN 2 RELATED, LARGE GTPASE FAMILY. GMPPNP, GPPNHP.
310	1osm	A	9	99	1.5e-15	1.73	-0.20		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
310	1qq4	A	14	119	1.2e-11	1.84	0.04		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
310	1qq4	A	8	96	3e-09	1.29	-0.08		ALPHA-LYTIC PROTEASE; CHAIN: A;	HYDROLASE DOUBLE BETA BARREL, BACTERIAL SERINE PROTEASE
310	1tal		14	119	1e-11	1.63	-0.06		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
310	1tal		8	99	1.2e-10	1.03	-0.20		ALPHA-LYTIC PROTEASE; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, LOW TEMPERATURE, HYDROLASE, 2 SERINE PROTEINASE
311	1a1h	A	116	195	8.5e-18	-0.02	0.27		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
311	1a1h	A	339	448	1.2e-39	-0.24	0.11		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
311	1a1h	A	619	728	6e-37	-0.46	0.12		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
311	1mey	C	105	195	3.4e-32	-0.06	0.22		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	142	223	1.7e-39	-0.08	0.95		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	170	251	1.7e-42	0.19	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	198	279	1.2e-44	0.17	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	226	307	3.4e-46	0.27	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	254	335	1.7e-46	-0.02	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	282	363	8.5e-47	-0.26	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	310	391	1.5e-46	-0.08	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	338	419	1.7e-46	0.07	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	366	447	3.4e-47	0.08	1.00		DNA: CHAIN: A, B, D, E: CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(ZINC FINGER/DNA)
311	1mey	C	394	475	6.8e-49	0.32	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	422	503	1e-49	0.06	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	450	531	3.4e-49	0.18	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	478	559	1.2e-48	0.37	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	478	560	3.4e-49			108.14	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	506	587	8.5e-49	0.14	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	534	615	1.5e-48	0.14	0.99		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	562	643	6.8e-49	-0.00	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	590	671	6.8e-49	0.04	0.82		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	618	699	1.7e-49	-0.22	0.94		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	646	727	1.2e-50	-0.03	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	674	755	1.2e-50	0.23	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	702	783	6.8e-51	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	730	811	3.4e-50	0.08	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	758	839	1.7e-50	-0.03	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1mey	C	786	852	1.5e-40	-0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
311	1tf6	A	199	345	1.7e-34	0.00	0.98		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	394	560	1.7e-37			116.05	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	395	540	1.7e-37	0.21	0.88		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	507	652	3.4e-36	-0.03	0.48		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	563	708	1.4e-36	-0.36	0.45		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	619	764	1.4e-36	-0.22	0.46		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1tf6	A	703	852	6.8e-38	0.19	0.98		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
311	1ubd	C	116	223	5.1e-25	-0.02	0.86		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	147	251	1.5e-41	-0.11	0.94		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	168	279	6e-51	-0.09	0.66		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	201	307	8.5e-32	0.06	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	203	307	6e-53	-0.15	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	224	336	3e-52	-0.10	0.72		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
311	1ubd	C	308	447	1.5e-48	-0.40	0.54		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	371	476	3e-50	-0.17	0.86		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	374	475	1e-34	-0.10	0.80		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	394	503	1.5e-54	-0.13	0.69		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	420	559	7.5e-57	-0.17	0.46		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	430	531	1.7e-34	0.22	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, B;	INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	458	559	1.7e-33	-0.07	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	504	615	1.3e-51	0.16	0.80		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	598	699	1.7e-34	-0.42	0.21		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	616	755	3e-49	-0.41	0.21		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	626	727	1.7e-34	-0.33	0.45		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	672	784	3e-57			93.95	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	682	783	1.7e-34	-0.02	0.89		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	700	811	3e-57	-0.12	0.90		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	728	839	1.5e-54	0.00	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	738	839	1.4e-34	-0.08	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	1ubd	C	756	852	1.2e-44	-0.20	0.51		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
311	2gli	A	119	250	1.4e-28	-0.17	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	143	281	6e-55	0.02	0.53		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	198	334	8.5e-32	-0.17	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	201	337	1.5e-65	0.34	0.86		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	226	365	4.5e-66	0.06	0.95		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	310	477	4.5e-64	0.04	0.60		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	346	474	5.1e-32	0.08	0.84		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	394	533	7.5e-72	0.06	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	422	561	7.5e-72			102.20	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										BINDING PROTEIN/DNA)
311	2gli	A	430	558	3.4e-33	0.32	0.78		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	478	617	1.2e-68	-0.09	0.65		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	570	698	1.2e-33	-0.19	0.07		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	654	782	1.2e-33	0.17	0.55		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	674	841	4.5e-70	-0.08	0.63		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	682	810	5.1e-33	0.03	0.62		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	710	841	6.8e-34	-0.14	0.84		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	730	852	6e-60	0.12	0.80		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
311	2gli	A	738	851	3.4e-29	0.17	0.80		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
312	1b34	A	7	81	4.5e-23	0.29	0.30		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1; CHAIN: A; SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2;	RNA BINDING PROTEIN SNRNP, SPLICING, SPLICEOSOME, SM, CORE SNRNP DOMAIN, 2 SYSTEMIC LUPUS ERYTHEMATOSUS, SLE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: B;	
312	1b34	A	9	71	3.4e-17	0.08	0.04		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1; CHAIN: A; SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2; CHAIN: B;	RNA BINDING PROTEIN SNRNP, SPLICING, SPLICEOSOME, SM, CORE SNRNP DOMAIN, 2 SYSTEMIC LUPUS ERYTHEMATOSUS, SLE
312	1b34	B	7	72	1.2e-12	0.46	0.18		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D1; CHAIN: A; SMALL NUCLEAR RIBONUCLEOPROTEIN SM D2; CHAIN: B;	RNA BINDING PROTEIN SNRNP, SPLICING, SPLICEOSOME, SM, CORE SNRNP DOMAIN, 2 SYSTEMIC LUPUS ERYTHEMATOSUS, SLE
312	1d3b	A	5	72	1.7e-14	0.28	0.07		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3; CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
312	1d3b	A	7	76	1.1e-20	0.63	0.05		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3; CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
312	1d3b	B	10	78	7.5e-18	0.34	-0.06		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3; CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
312	1d3b	B	9	70	1.7e-15	-0.01	0.04		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3; CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
312	1d3b	D	4	70	6.8e-16	0.76	-0.05		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3; CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
312	1d3b	D	9	78	1.5e-17	0.11	-0.01		SMALL NUCLEAR RIBONUCLEOPROTEIN SM D3;	RNA BINDING PROTEIN D3 CORE SNRNP PROTEIN; B CORE SNRNP

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, C, E, G, I, K; SMALL NUCLEAR RIBONUCLEOPROTEIN ASSOCIATED CHAIN: B, D, F, H, J, L;	PROTEIN SNRNP, SPLICING, SM, CORE SNRNP DOMAIN, SYSTEMIC LUPUS 2 ERYTHEMATOSUS, SLE, RNA BINDING PROTEIN
314	1b8q	A	101	173	1.3e-06	0.08	0.15		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
314	1be9	A	113	175	8.5e-05	0.19	0.99		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
314	1pdr		113	175	0.0012	0.13	0.90		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
314	1qlc	A	119	172	0.00034	0.27	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
314	3pdz	A	109	190	3e-09	0.73	0.76		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
316	1a4y	A	805	893	4.5e-05	0.52	0.60		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
316	1a4y	A	815	877	7.5e-08	0.54	1.00		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
316	1aoj	A	590	647	6e-16	-0.80	0.30		EPS8; CHAIN: A, B;	SIGNAL TRANSDUCTION SRC HOMOLOGY DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, EPS8, PROLINE RICH PEPTIDE
316	1tud		577	627	1.1e-07	-0.30	0.64		ALPHA-SPECTRIN; CHAIN: NULL;	CYTOSKELETON CAPPING PROTEIN, CALCIUM-BINDING, DUPLICATION, REPEAT, 2 SH3 DOMAIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										CYTOSKELETON
316	2nmb	A	155	263	9e-14	-0.13	0.10		NUMB PROTEIN; CHAIN: A; GPPY PEPTIDE; CHAIN: B;	CELL CYCLE/GENE REGULATION COMPLEX, SIGNAL TRANSDUCTION, PHOSPHOTYROSINE BINDING 2 DOMAIN (PTB), ASYMETR IC CELL DIVISION, CELL CYCLE/GENE 3 REGULATION
318	1a5f	H	38	246	1.4e-20			69.16	MONOCLONAL ANTI-E-SELECTIN 7A9 ANTIBODY; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB, ANTIBODY, ANTI-E-SELECTIN
318	1adq	L	41	241	6.8e-29	0.18	0.35		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
318	1ae6	H	38	255	1.7e-22			64.20	ANTIBODY CTM01; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB FRAGMENT, HUMANISATION
318	1afv	H	38	236	1.7e-23	0.11	1.00		HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 CAPSID CHAIN: A, B; ANTIBODY FAB25.3 FRAGMENT; CHAIN: H, K, L, M;	COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN) HIV-1 CA, HIV CA, HIV P24, P24; FAB, FAB LIGHT CHAIN, FAB HEAVY CHAIN COMPLEX (VIRAL CAPSID/IMMUNOGLOBULIN), HIV, CAPSID PROTEIN, 2 P24
318	1aqk	H	39	247	3.4e-20			65.54	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
318	1aqk	L	40	260	1.7e-26			64.54	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
318	1aqk	L	41	241	1.7e-26	0.21	0.74		FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
318	1ayl	H	50	236	8.5e-23	-0.10	0.28		TP7 FAB; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, ANTIBODY, FAB, ENZYME INHIBITOR, PCR, 2 HOT START
318	1b2w	H	39	247	1.2e-19			63.67	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRYCTURE, GAMMA- 3 INTERFERON, IMMUNE SYSTEM
318	1b2w	L	38	259	1.5e-23			64.13	ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRYCTURE, GAMMA- 3 INTERFERON, IMMUNE SYSTEM
318	1b2w	L	39	232	1.5e-23	-0.00	0.07		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRYCTURE, GAMMA- 3 INTERFERON, IMMUNE SYSTEM
318	1b4j	H	39	247	1.2e-19			67.97	ANTIBODY; CHAIN: L, H;	ANTIBODY ENGINEERING ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODIES, 2 FAB, X-RAY STRUCTURES, GAMMA-INTERFERON
318	1baf	H	37	259	8.5e-21			65.16	IMMUNOGLOBULIN FAB FRAGMENT OF MURINE MONOCLONAL ANTIBODY AN02 COMPLEX 1BAF 3 WITH ITS HAPTEN (2,2,6,6-TETRAMETHYL-1-PIPERIDINYLOXY- 1BAF 4 DINITROPHENYL) 1BAF 5	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
318	1bih	A	90	246	8.5e-20	0.30	0.41		HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
318	1bj1	L	39	232	1e-22	0.22	0.11		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
318	1bjm	A	40	241	1.7e-26	0.07	0.11		LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENCE-JONES PROTEIN; 1BJM 8 BENCE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
318	1bm3	H	37	259	6.8e-21			68.28	IMMUNOGLOBULIN OPG2 FAB, CONSTANT DOMAIN; CHAIN: L; IMMUNOGLOBULIN OPG2 FAB, VARIABLE DOMAIN; CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN
318	1cf8	H	37	254	5.1e-22			64.09	CATALYTIC ANTIBODY 19A4 (LIGHT CHAIN); CHAIN: L; CATALYTIC ANTIBODY 19A4 (HEAVY CHAIN); CHAIN: H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, TERPENOID SYNTHASE, CARBOCATION, 2 CYCLIZATION CASCADE
318	1cic	B	38	257	5.1e-22			67.07	IG HEAVY CHAIN V REGIONS; CHAIN: A; IG HEAVY CHAIN V REGIONS; CHAIN: B; IG HEAVY CHAIN V REGIONS; CHAIN: C; IG HEAVY CHAIN V REGIONS; CHAIN: D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, FAB COMPLEX, IDIOTOPE, ANTI-IDIOTOPE
318	1cs6	A	45	247	1.7e-32	0.03	0.77		AXONIN-1; CHAIN: A;	CELL ADHESION NEURAL CELL ADHESION
318	1ct8	B	38	259	3.4e-22			64.34	7C8 FAB FRAGMENT; SHORT CHAIN; CHAIN: A, C; 7C8 FAB FRAGMENT; LONG CHAIN; CHAIN: B, D	IMMUNE SYSTEM ABZYME TRANSITION STATE ANALOG, IMMUNE SYSTEM
318	1cvs	C	174	249	1.7e-12	0.27	0.34		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
318	1cvs	D	174	249	1.7e-12	0.22	0.34		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										RECEPTOR
318	1dee	A	39	232	3.4e-23	0.15	0.06		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
318	1ev2	E	173	249	1.7e-13	0.13	0.25		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2: CHAIN: E. F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
318	1evt	C	174	249	1.7e-12	0.37	0.13		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
318	1fai	H	38	254	3.4e-19			63.84	IMMUNOGLOBULIN FAB FRAGMENT FROM A MONOCLONAL ANTI-ARSONATE ANTIBODY, R19.9 1FAI 3 (IGG2B,KAPPA) 1FAI 4	
318	1fhg	A	154	247	1.5e-08	0.27	0.16		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
318	1fvd	B	37	247	5.1e-21			66.24	IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
318	1hnf		43	232	1.3e-23	0.02	0.10		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (HUMAN) 1HNF 3	
318	1iai	H	38	254	5.1e-20			65.01	IDIOTYPIC FAB 730.1.4 (IGG1) OF VIRUS 1IAI 5 CHAIN: L, H: 1IAI 7 ANTI-IDIOTYPIC FAB 409.5.3 (IGG2A); 1IAI 9 CHAIN: M, I 1IAI 10	COMPLEX (IMMUNOGLOBULIN IGG1/IGG2A)
318	1lil	A	40	241	1.7e-25	0.18	0.13		LAMBDA III BENCE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
318	1nca	H	38	254	1.5e-21			67.34	HYDROLASE(O-GLYCOSYL) N9 NEURAMINIDASE-NC41 (E.C.3.2.1.18) COMPLEX WITH FAB 1NCA 3	
318	1nsn	H	37	254	1.4e-22			65.27	IGG FAB (IGG1, KAPPA); 1NSN 4 CHAIN: L, H; 1NSN 5 STAPHYLOCOCCAL NUCLEASE; 1NSN 9 CHAIN: S; 1NSN 10	COMPLEX (IMMUNOGLOBULIN/HYDROLASE) N10 FAB IMMUNOGLOBULIN; 1NSN 7 STAPHYLOCOCCAL NUCLEASE RIBONUCLEATE, 1NSN 11 IMMUNOGLOBULIN, STAPHYLOCOCCAL NUCLEASE 1NSN 25
318	1wio	A	47	262	7.5e-28	0.19	0.29		T-CELL SURFACE GLYCOPROTEIN CD4; CHAIN: A, B;	GLYCOPROTEIN CD4; IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, T-CELL, 2 MHC LIPOPROTEIN, POLYMORPHISM
318	25c8	H	38	255	5.1e-23			64.25	IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
318	25c8	H	50	236	5.1e-23	0.12	0.19		IGG 5C8; CHAIN: L, H;	CATALYTIC ANTIBODY CATALYTIC ANTIBODY, FAB, RING CLOSURE REACTION
318	2cgr	H	37	254	1.2e-17			65.11	IMMUNOGLOBULIN IGG2B (KAPPA) FAB FRAGMENT COMPLEXED WITH ANTIGEN 2CGR 3 N-(P-CYANOPHENYL)-N'-(DIPHENYLEMETHYL) GUANIDINEACETIC ACID 2CGR 4	
318	2fb4	L	40	241	1.5e-25	0.29	0.37		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
318	2fgw	L	39	232	1.2e-23	0.27	0.01		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
318	2mcg	l	40	241	1.2e-27	0.14	0.30		IMMUNOGLOBULIN IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER (/MCG$) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
318	2pcp	B	38	255	3.4e-21			68.25	IMMUNOGLOBULIN; CHAIN: A, B,	IMMUNOGLOBULIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									C, D;	IMMUNOGLOBULIN
318	32c2	B	50	236	3.4e-23	0.21	0.13		IGG1 ANTIBODY 32C2; CHAIN: A; IGG1 ANTIBODY 32C2; CHAIN: B;	IMMUNE SYSTEM FAB, ANTIBODY, AROMATASE, P450
318	3fct	B	37	247	8.5e-19			66.99	METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: A, C: METAL CHELATASE CATALYTIC ANTIBODY; CHAIN: B, D;	IMMUNE SYSTEM METAL CHELATASE, CATALYTIC ANTIBODY, FAB FRAGMENT, IMMUNE 2 SYSTEM
318	3ncm	A	168	245	4.5e-09	0.09	-0.14		NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM; CHAIN: A;	CELL ADHESION PROTEIN NCAM MODULE 2; CELL ADHESION, GLYCOPROTEIN, HEPARIN-BINDING, GPI-ANCHOR, 2 NEURAL ADHESION MOLECULE, IMMUNOGLOBULIN FOLD, HOMOPHILIC 3 BINDING, CELL ADHESION PROTEIN
318	7fab	L	40	241	1.7e-26	0.00	0.17		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
318	8fab	A	43	241	1.4e-26	0.39	0.18		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
319	1c1y	A	8	171	1.4e-63			109.08	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
319	1c1y	A	9	171	1.4e-63	0.82	1.00		RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
319	1ctq	A	8	172	6.8e-65			108.57	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
319	1ctq	A	9	171	6.8e-65	0.88	1.00		TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
319	1cxz	A	3	172	3.4e-55			108.33	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
319	1d5c	A	9	165	6e-67	0.83	1.00		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING
319	1d5c	A	9	169	5.1e-63	0.87	1.00		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G-PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING
319	1ds6	A	9	170	1.5e-55	0.73	1.00		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 2; CHAIN: A; RHO GDP-DISSOCIATION INHIBITOR 2; CHAIN: B;	SIGNALING PROTEIN P21-RAC2; RHO GDI 2, RHO-GDI BETA, LY-GDI; BETA SANDWHICH, PROTEIN-PROTEIN COMPLEX, G-DOMAIN, 2 IMMUNOGLOBULIN FOLD, WALKER FOLD, GTP-BINDING PROTEIN
319	1ek0	A	9	169	5.1e-61	0.93	1.00		GTP-BINDING PROTEIN YPT51; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G PROTEIN, VESICULAR TRAFFIC, GTP HYDROLYSIS, YPT/RAB 2 PROTEIN, ENDOCYTOSIS, HYDROLASE
319	1kao		8	172	1.2e-59			109.78	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
319	1kao		9	169	1.2e-59	0.96	1.00		RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
319	1tx4	B	6	170	1.1e-56			97.67	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
319	1tx4	B	7	170	1.1e-56	0.65	1.00		P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
319	1zbd	A	3	178	5.1e-70			154.65	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
319	1zbd	A	5	175	5.1e-70	0.92	1.00		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
319	3rab	A	4	172	1.5e-70	0.71	1.00		RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
319	3rab	A	4	172	1.5e-70			170.48	RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
321	1b0x	A	227	287	1.5e-05	1.26	0.99		EPHA4 RECEPTOR TYROSINE KINASE; CHAIN: A;	TRANSFERASE RECEPTOR TYROSINE KINASE, PROTEIN INTERACTION MODULE, 2 DIMERIZATION DOMAIN, TRANSFERASE
321	1b4f	A	226	297	1.2e-13	0.85	0.74		EPHB2; CHAIN: A, B, C, D, E, F, G, H;	SIGNAL TRANSDUCTION SAM DOMAIN, EPH RECEPTOR, SIGNAL TRANSDUCTION, OLIGOMER
321	1sgg		226	287	3e-06	0.84	0.92		EPHRIN TYPE-B RECEPTOR 2; CHAIN: NULL;	TYROSINE-PROTEIN KINASE NMR, RECEPTOR OLIGOMERIZATION, EPH RECEPTORS, TYROSINE 2 PHOSPHORYLATION, SIGNAL TRANSDUCTION, TYROSINE-PROTEIN 3 KINASE
323	1a17		114	266	3.4e-12	0.15	0.43		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1a17		130	279	4.5e-14	0.30	-0.01		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1a17		157	318	6e-08	0.17	-0.02		SERINE/THREONINE PROTEIN	HYDROLASE TETRATRICOPEPTIDE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PHOSPHATASE 5; CHAIN: NULL;	TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1a17		246	380	6.8e-13	0.22	0.22		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1a17		293	400	1.7e-13	0.34	-0.12		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1a17		4	143	5.1e-16	0.43	0.07		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
323	1b89	A	11	275	0.00017	0.05	0.04		CLATHRIN HEAVY CHAIN; CHAIN: A;	CLATHRIN CLATHRIN, TRISKELION, COATED VESICLES, ENDOCYTOSIS, SELF- 2 ASSEMBLY, ALPHA-ALPHA SUPERHELIX
323	1e96	B	162	318	6.8e-11	0.31	0.11		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1; CHAIN: A; NEUTROPHIL CYTOSOL FACTOR 2 (NCF-2) CHAIN: B;	SIGNALLING COMPLEX RAC1; P67PHOX; SIGNALLING COMPLEX, GTPASE, NADPH OXIDASE, PROTEIN-PROTEIN 2 COMPLEX, TPR MOTIF
323	1e96	B	2	109	6.8e-10	0.31	-0.06		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1; CHAIN: A; NEUTROPHIL CYTOSOL FACTOR 2 (NCF-2) CHAIN: B;	SIGNALLING COMPLEX RAC1; P67PHOX; SIGNALLING COMPLEX, GTPASE, NADPH OXIDASE, PROTEIN-PROTEIN 2 COMPLEX, TPR MOTIF
323	1e96	B	245	392	1.2e-08	0.16	-0.14		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1; CHAIN: A; NEUTROPHIL CYTOSOL FACTOR 2 (NCF-2) CHAIN: B;	SIGNALLING COMPLEX RAC1; P67PHOX; SIGNALLING COMPLEX, GTPASE, NADPH OXIDASE, PROTEIN-PROTEIN 2 COMPLEX, TPR MOTIF
323	1e96	B	82	232	1.2e-10	0.27	-0.02		RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1; CHAIN: A; NEUTROPHIL CYTOSOL FACTOR 2 (NCF-2) CHAIN: B;	SIGNALLING COMPLEX RAC1; P67PHOX; SIGNALLING COMPLEX, GTPASE, NADPH OXIDASE, PROTEIN-PROTEIN 2 COMPLEX, TPR MOTIF
323	1elr	A	11	114	1.7e-15	0.50	0.90		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
323	1elr	A	121	233	1.2e-12	0.42	0.22		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	169	274	3.4e-13	0.04	0.06		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	1	74	1e-09	0.40	-0.01		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	212	313	1.2e-15	0.58	-0.05		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	252	356	1.2e-13	0.05	0.05		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	332	411	1e-11	0.04	-0.18		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	56	157	1.5e-07	-0.03	0.21		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elr	A	88	194	1.7e-13	0.19	0.28		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
323	1elw	A	126	244	3.4e-11	0.18	0.81		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
323	1elw	A	249	366	1e-11	0.20	0.19		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D:	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
323	1elw	A	293	393	3.4e-11	0.29	-0.08		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
323	1elw	A	4	121	3.4e-14	0.56	0.62		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
323	1elw	A	81	208	1.2e-08	0.18	-0.11		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
323	1fch	A	104	410	1e-31	0.02	-0.02		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
323	1fch	A	11	317	1.2e-29	0.37	0.87		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
323	1fch	A	2	263	3.4e-23	0.36	0.99		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
323	1qqe	A	120	375	3.4e-10	0.14	0.58		VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
323	1qqe	A	221	388	3.4e-10	0.01	-0.09		VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
323	1qqe	A	3	188	1e-11	0.48	0.19		VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
323	1qqe	A	68	359	3.4e-10			54.55	VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
324	1b4f	A	28	74	0.00045	0.19	0.90		EPHB2; CHAIN: A, B, C, D, E, F, G, H;	SIGNAL TRANSDUCTION SAM DOMAIN, EPH RECEPTOR, SIGNAL TRANSDUCTION, OLIGOMER
329	1b7f	A	421	559	5.1e-20	-0.15	0.98		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP *UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
329	1cvj	A	423	547	1.7e-21	0.00	0.52		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	
329	1cvj	B	423	535	3.4e-20	0.09	0.63		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
329	1cvj	F	423	502	3.4e-17	0.54	0.87		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
329	1cvj	H	423	535	1.4e-17	-0.03	0.49		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
329	1d8z	A	418	496	6.8e-19	0.12	0.82		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
329	1hal		416	544	1.7e-17	-0.01	0.03		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
329	2sxl		421	496	1.2e-16	0.37	0.96		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
329	2upl	A	415	550	1e-17	-0.04	0.05		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
329	3sxl	A	421	559	1.7e-19	0.05	0.89		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										DOSAGE COMPENSATION
332	1adq	L	57	268	6e-98	0.86	1.00		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
332	1adq	L	57	268	6e-98			301.73	IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
332	1aqk	L	56	268	6.8e-88			318.27	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
332	1b2w	L	55	267	5.1e-90	0.76	1.00		ANTIBODY (LIGHT CHAIN); CHAIN: L; ANTIBODY (HEAVY CHAIN); CHAIN: H;	IMMUNE SYSTEM IMMUNOGLOBULIN; IMMUNOGLOBULIN ANTIBODY ENGINEERING, HUMANIZED AND CHIMERIC ANTIBODY, FAB, 2 X-RAY STRUCTURE, THREE-DIMENSIONAL STRYCTURE, GAMMA- 3 INTERFERON, IMMUNE SYSTEM
332	1bjm	A	55	268	3.4e-85			322.11	LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENCE-JONES PROTEIN; 1BJM 8 BENCE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
332	1bwm	A	7	161	3.4e-21	-0.07	0.33		ALPHA-BETA T CELL RECEPTOR (TCR) (D10); CHAIN: A;	IMMUNE SYSTEM IMMUNOGLOBULIN, IMMUNORECEPTOR, IMMUNE SYSTEM
332	1dee	A	55	267	1e-90	0.84	1.00		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
332	1dzb	A	1	162	5.1e-60	0.09	0.46		SCFV FRAGMENT 1F9; CHAIN: A, B; TURKEY EGG-WHITE LYSOZYME C; CHAIN: X, Y;	COMPLEX (ANTIBODY ANTIGEN) 1,4-BETA-N-ACETYLMURAMIDASE C; SINGLE-DOMAIN ANTIBODY, TURKEY EGG-WHITE LYSOZYME, 2 ANTIBODY-PROTEIN COMPLEX, SINGLE-CHAIN FV FRAGMENT
332	1f3r	B	1	164	1.4e-61	0.14	0.98		ACETYLCHOLINE RECEPTOR ALPHA; CHAIN: A; FV ANTIBODY FRAGMENT; CHAIN: B;	IMMUNE SYSTEM IG-FOLD, IMMUNO COMPLEX, ANTIBODY-ANTIGEN, BETA-TURN
332	1igt	A	55	267	1.2e-89	0.68	1.00		IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
332	1lil	A	57	268	4.5e-99	0.86	1.00		LAMBDA III BENCE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN
332	1lil	A	58	268	4.5e-99			299.68	LAMBDA III BENCE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN
332	1lmk	A	1	162	3.4e-59	0.12	0.92		IMMUNOGLOBULIN ANTI-PHOSPHATIDYLINOSITOL SPECIFIC PHOSPHOLIPASE C DIABODY 1LMK 3 SYNONYMS: L5MK16 DIABODY, SINGLE-CHAIN FV DIMER 1LMK 4	
332	1mcp	L	55	267	3.4e-91	0.79	1.00		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB FRAGMENT (MC/PC$603) 1MCP 4	
332	1mcp	L	55	267	3.4e-91			202.00	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB FRAGMENT (MC/PC$603) 1MCP 4	
332	1mcw	W	55	268	1e-82			294.22	IMMUNOGLOBULIN IMMUNOGLOBULIN HETEROLOGOUS LIGHT CHAIN DIMER 1MCW 3 (/MCG$-/WEIR$ HYBRID) 1MCW 4	
332	1mfa		1	161	3.4e-21	-0.34	0.01		IMMUNOGLOBULIN FV FRAGMENT (MURINE SE155-4) COMPLEX WITH THE TRISACCHARIDE: 1MFA 3 ALPHA-D-GALACTOSE(1-2)[ALPHA-D-ABEQUOSE(1-3)]ALPHA- 1MFA 4 D-MANNOSE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									(P1-OME) (PART OF THE CELL-SURFACE CARBOHYDRATE 1MFA 5 OF PATHOGENIC SALMONELLA) 1MFA 6	
332	1nca	L	55	267	5.1e-91	0.78	1.00		HYDROLASE(O-GLYCOSYL) N9 NEURAMINIDASE-NC41 (E.C.3.2.1.18) COMPLEX WITH FAB 1NCA 3	
332	1nqb	A	1	163	5.1e-61	0.17	0.53		SINGLE-CHAIN ANTIBODY FRAGMENT; CHAIN: A, C;	IMMUNOGLOBULIN VARIABLE HEAVY (VH) DOMAIN, VARIABLE LIGHT (VL) ANTIBODY FRAGMENT, MULTIVALENT ANTIBODY, DIABODY, DOMAIN 2 SWAPPING, IMMUNOGLOBULIN
332	1qlr	A	55	267	1.5e-89	0.65	1.00		IGM KAPPA CHAIN V-III (KAU COLD AGGLUTININ); CHAIN: A, C; IGM FAB REGION IV-J(H4)-C (KAU COLD AGGLUTININ); CHAIN: B, D;	IMMUNOGLOBULIN IMMUNOGLOBULIN, AUTOANTIBODY, COLD AGGLUTININ, HUMAN IGM 2 FAB FRAGMENT
332	1qok	A	1	162	1.7e-61	0.45	0.42		MFE-23 RECOMBINANT ANTIBODY FRAGMENT; CHAIN: A;	IMMUNOGLOBULIN IMMUNOGLOBULIN, SINGLE-CHAIN FV, ANTI-CARCINOEMBRYONIC 2 ANTIGEN
332	1sbs	L	55	267	3.4e-92	0.89	1.00		MONOCLONAL ANTIBODY 3A2; CHAIN: H, L;	MONOCLONAL ANTIBODY MONOCLONAL ANTIBODY, FAB-FRAGMENT, REPRODUCTION
332	2fb4	L	55	268	6.8e-87			326.11	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
332	2mcg	l	55	268	1.7e-86			304.84	IMMUNOGLOBULIN IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER (/MCG$) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
332	7fab	L	55	264	3e-95			290.47	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
332	7fab	L	56	264	3e-95	0.85	1.00		IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
332	8fab	A	58	264	5.1e-87			291.96	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									(LAMBDA, HIL) 8FAB 3	
338	1fl3	L	39	117	0.00034	-0.04	0.22		BLUE FLUORESCENT ANTIBODY (19G2)-HEAVY CHAIN; CHAIN: H, A; BLUE FLUORESCENT ANTIBODY (19G2)-LIGHT CHAIN; CHAIN: L, B;	IMMUNE SYSTEM IMMUNOGLOBULIN FOLD
342	1ekl	A	225	349	3.4e-14	-0.04	0.19		EPOXIDE HYDROLASE; CHAIN: A, B;	HYDROLASE HOMODIMER, ALPHA/BETA HYDROLASE FOLD, DISUBSTITUTED UREA 2 INHIBITOR
342	1ekl	B	132	349	1.5e-17	0.25	0.54		EPOXIDE HYDROLASE; CHAIN: A, B;	HYDROLASE HOMODIMER, ALPHA/BETA HYDROLASE FOLD, DISUBSTITUTED UREA 2 INHIBITOR
342	1fez	A	130	330	4.5e-29	0.37	0.82		PHOSPHONOACETALDEHYDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HAD-FAMILY ALPHA/BETA CORE DOMAIN, MG(II) BINDING SITE, 5- 2 HELIX BUNDLE
342	1fez	A	130	366	1.5e-23	0.56	1.00		PHOSPHONOACETALDEHYDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HAD-FAMILY ALPHA/BETA CORE DOMAIN, MG(II) BINDING SITE, 5- 2 HELIX BUNDLE
342	1qq5	A	130	386	3.4e-26			51.58	L-2-HALOACID DEHALOGENASE; CHAIN: A, B;	HYDROLASE L-2-HALOACID DEHALOGENASE, HYDROLASE
342	1qq5	A	131	362	3.4e-26	0.32	0.65		L-2-HALOACID DEHALOGENASE; CHAIN: A, B;	HYDROLASE L-2-HALOACID DEHALOGENASE, HYDROLASE
342	1zrn		130	362	1.7e-28			57.26	L-2-HALOACID DEHALOGENASE; CHAIN: NULL;	DEHALOGENASE DEHALOGENASE, HYDROLASE
342	1zrn		131	361	1.7e-28	0.29	0.76		L-2-HALOACID DEHALOGENASE; CHAIN: NULL;	DEHALOGENASE DEHALOGENASE, HYDROLASE
343	1a1h	A	129	213	8.5e-24	0.05	-0.05		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
343	1a1h	A	161	241	3.4e-30	0.13	0.12		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
343	1mey	C	145	213	3.4e-38	-0.21	0.10		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN: CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
343	1mey	C	160	241	6.8e-50	0.09	0.54		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	188	269	5.1e-50	-0.08	0.89		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	216	297	5.1e-50	0.20	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	244	325	3.4e-50	0.22	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	272	353	1.4e-49	0.47	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	272	354	3.4e-50			103.55	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	300	357	3.4e-33	0.42	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	C	39	142	5.1e-43	-0.12	0.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	G	158	185	1.2e-12	0.50	0.71		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1mey	G	37	64	1.7e-11	-0.39	0.13		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
343	1tf6	A	161	313	8.5e-38	-0.20	0.66		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
343	1tf6	A	187	353	8.5e-38			89.34	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
343	1tf6	A	217	355	3.4e-35	0.13	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
343	1ubd	C	168	269	5.1e-35	-0.19	0.69		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
343	1ubd	C	214	325	1.2e-52	-0.09	0.93		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
343	1ubd	C	242	353	6e-53	0.03	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
343	1ubd	C	244	354	6e-53			86.36	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
343	1ubd	C	252	353	6.8e-34	0.09	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
343	2gli	A	157	268	1.2e-31	0.00	0.27		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
343	2gli	A	188	327	1.2e-61	0.41	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
343	2gli	A	216	353	1.5e-67	0.42	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
343	2gli	A	216	355	1.5e-67			95.61	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
343	2gli	A	224	352	3.4e-33	0.43	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
343	2gli	A	40	243	3e-23	-0.10	0.00		ZINC FINGER PROTEIN GLI1;	COMPLEX (DNA-BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A: DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
345	1bbz	A	7	63	4.5e-15	-0.10	0.72		ABL TYROSINE KINASE; CHAIN: A, C, E, G; PEPTIDE P41; CHAIN: B, D, F, H;	COMPLEX (TRANSFERASE/PEPTIDE) COMPLEX (TRANSFERASE/PEPTIDE), SIGNAL TRANSDUCTION, 2 SH3 DOMAIN
345	1gbq	A	8	63	3e-16	-0.22	0.88		GRB2; CHAIN: A; SOS-1; CHAIN: B;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), SH3 DOMAIN
345	1gbr	A	8	65	3e-16	-0.04	0.98		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2, N-TERMINAL 1GBR 3 SH3 DOMAIN) COMPLEXED WITH SOS-A PEPTIDE 1GBR 4 (NMR, 29 STRUCTURES) 1GBR 5	
345	1gfc		8	63	3e-15	0.27	0.89		ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
345	1pht		8	71	1.2e-15	-0.32	0.33		PHOSPHATIDYLINOSITOL 3-KINASE P85-ALPHA SUBUNIT; 1PHT 6 CHAIN: NULL; 1PHT 7	PHOSPHOTRANSFERASE PI3K SH3; 1PHT 9 PHOSPHATIDYLINOSITOL 3-KINASE, P85-ALPHA SUBUNIT, SH3 DOMAIN 1PHT 21
345	1pks		8	63	1.5e-14	-0.24	0.30		PHOSPHOTRANSFERASE PHOSPHATIDYLINOSITOL 3-KINASE (E.C.2.7.1.137) (PI3K) 1PKS 3 (SH3 DOMAIN) (NMR, MINIMIZED AVERAGE STRUCTURE) 1PKS 4	
345	1pwt		1	63	7.5e-16	-0.09	0.99		ALPHA SPECTRIN; CHAIN: NULL;	CIRCULAR PERMUTANT PWT; CIRCULAR PERMUTANT, SH3 DOMAIN, CYTOSKELETON
345	1qkw	A	8	63	7.5e-16	0.13	0.98		ALPHA II SPECTRIN; CHAIN: A;	CYTOSKELETON CYTOSKELETON, MEMBRANE, SH3 DOMAIN
345	1sem	A	8	58	6e-15	0.30	0.92		SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
348	2occ	K	30	78	8.5e-27	-0.76	0.60		CYTOCHROME C OXIDASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	OXIDOREDUCTASE FERROCYTOCHROME C\:OXYGEN OXIDOREDUCTASE; OXIDOREDUCTASE, CYTOCHROME(C)-OXYGEN, CYTOCHROME C 2 OXIDASE
348	2occ	K	30	78	8.5e-27			69.07	CYTOCHROME C OXIDASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	OXIDOREDUCTASE FERROCYTOCHROME C\:OXYGEN OXIDOREDUCTASE; OXIDOREDUCTASE, CYTOCHROME(C)-OXYGEN, CYTOCHROME C 2 OXIDASE
355	1bxe	A	66	175	5.1e-43	0.90	1.00		RIBOSOMAL PROTEIN L22; CHAIN: A;	RNA BINDING PROTEIN RIBOSOMAL PROTEIN, PROTEIN SYNTHESIS, RNA BINDING, 2 ANTIBIOTICS RESISTANCE, RNA BINDING PROTEIN
355	1ffk	O	54	174	3.4e-23	0.21	0.60		23S RRNA; CHAIN: 0; 5S RRNA; CHAIN: 9; RIBOSOMAL PROTEIN L2; CHAIN: A; RIBOSOMAL PROTEIN L3; CHAIN: B; RIBOSOMAL PROTEIN L4; CHAIN: C; RIBOSOMAL PROTEIN L5; CHAIN: D; RIBOSOMAL PROTEIN L7AE; CHAIN: E; RIBOSOMAL PROTEIN L10E; CHAIN: F; RIBOSOMAL PROTEIN L13; CHAIN: G; RIBOSOMAL PROTEIN L14; CHAIN: H; RIBOSOMAL PROTEIN L15E; CHAIN: I; RIBOSOMAL PROTEIN L15; CHAIN: J; RIBOSOMAL PROTEIN L18; CHAIN: K; RIBOSOMAL PROTEIN L18E; CHAIN: L; RIBOSOMAL PROTEIN L19; CHAIN: M; RIBOSOMAL PROTEIN L21E; CHAIN: N; RIBOSOMAL PROTEIN L22; CHAIN: O;	RIBOSOME 50S RIBOSOMAL PROTEIN L2P, HMAL2, HL4; 50S RIBOSOMAL PROTEIN L3P, HMAL3, HL1; 50S RIBOSOMAL PROTEIN L4E, HMAL4, HL6; 50S RIBOSOMAL PROTEIN L5P, HMAL5, HL13; 30S RIBOSOMAL PROTEIN HS6; 50S RIBOSOMAL PROTEIN L13P, HMAL13; 50S RIBOSOMAL PROTEIN L14P, HMAL14, HL27; 50S RIBOSOMAL PROTEIN L15P, HMAL15, HL9; 50S RIBOSOMAL PROTEIN L18P, HMAL18, HL12; 50S RIBOSOMAL PROTEIN L18E, HL29, L19; 50S RIBOSOMAL PROTEIN L19E, HMAL19, HL24; 50S RIBOSOMAL PROTEIN L21E, HL31; 50S RIBOSOMAL PROTEIN L22P, HMAL22, HL23; 50S RIBOSOMAL PROTEIN L23P, HMAL23, HL25, L21; 50S RIBOSOMAL PROTEIN L24P, HMAL24, HL16, HL15; 50S

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									RIBOSOMAL PROTEIN L23; CHAIN: P; RIBOSOMAL PROTEIN L24; CHAIN: Q; RIBOSOMAL PROTEIN L24E; CHAIN: R; RIBOSOMAL PROTEIN L29; CHAIN: S; RIBOSOMAL PROTEIN L30; CHAIN: T; RIBOSOMAL PROTEIN L31E; CHAIN: U; RIBOSOMAL PROTEIN L32E; CHAIN: V; RIBOSOMAL PROTEIN L37AE; CHAIN: W; RIBOSOMAL PROTEIN L37E; CHAIN: X; RIBOSOMAL PROTEIN L39E; CHAIN: Y; RIBOSOMAL PROTEIN L44E; CHAIN: Z; RIBOSOMAL PROTEIN L6; CHAIN: 1;	RIBOSOMAL PROTEIN L24E, HL21/HL22; 50S RIBOSOMAL PROTEIN L29P, HMAL29, HL33; 50S RIBOSOMAL PROTEIN L30P, HMAL30, HL20, HL16; 50S RIBOSOMAL PROTEIN L31E, L34, HL30; 50S RIBOSOMAL PROTEIN L32E, HL5; 50S RIBOSOMAL PROTEIN L37E, L35E; 50S RIBOSOMAL PROTEINS L39E, HL39E, HL46E; 50S RIBOSOMAL PROTEIN L44E, LA, HLA; 50S RIBOSOMAL PROTEIN L6P, HMAL6, HL10 RIBOSOME ASSEMBLY, RNA-RNA, PROTEIN-RNA, PROTEIN-PROTEIN
369	1d2h	A	70	190	1.2e-14	0.20	0.17		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE METHYLTRANSFERASE
369	2adm	A	66	209	6.8e-13	0.14	-0.11		ADENINE-N6-DNA-METHYLTRANSFERASE TAQI; CHAIN: A, B;	METHYLTRANSFERASE TRANSFERASE, METHYLTRANSFERASE, RESTRICTION SYSTEM
371	1a02	F	108	160	4.5e-13	-0.36	0.17		NFAT: CHAIN: N; C-FOS: CHAIN: F; C-JUN; CHAIN: J; DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR) NF-AT; TRANSCRIPTION FACTOR, PROTEIN-DNA COMPLEX, NFAT, NF-AT, 2 AP-1, FOS-JUN, QUATERNARY PROTEIN-DNA COMPLEX, CRYSTAL 3 STRUCTURE, TRANSCRIPTION SYNERGY, COMBINATORIAL GENE 4 REGULATION, COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR)
371	1a02	F	108	160	4.5e-13			62.39	NFAT; CHAIN: N; C-FOS; CHAIN: F; C-JUN; CHAIN: J; DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR) NF-AT; TRANSCRIPTION FACTOR, PROTEIN-DNA COMPLEX, NFAT, NF-AT, 2 AP-1, FOS-JUN, QUATERNARY PROTEIN-DNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										COMPLEX, CRYSTAL 3 STRUCTURE, TRANSCRIPTION SYNERGY, COMBINATORIAL GENE 4 REGULATION, COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR)
371	1a02	F	115	146	3.4e-10	-0.05	0.69		NFAT; CHAIN: N; C-FOS; CHAIN: F; C-JUN; CHAIN: J; DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR) NF-AT; TRANSCRIPTION FACTOR, PROTEIN-DNA COMPLEX, NFAT, NF-AT, 2 AP-1, FOS-JUN, QUATERNARY PROTEIN-DNA COMPLEX, CRYSTAL 3 STRUCTURE, TRANSCRIPTION SYNERGY, COMBINATORIAL GENE 4 REGULATION, COMPLEX (TRANSCRIPTION/NUCLEAR/NUCLEAR)
371	1fos	E	107	166	3.4e-10			70.24	COMPLEX (GENE-REGULATORY PROTEIN/DNA) C-JUN PROTO-ONCOGENE (TRANSCRIPTION FACTOR AP-1) DIMERIZED 1FOS 4 WITH C-FOS AND COMPLEXED WITH DNA 1FOS 5 COILED-COIL, DNA-BINDING PROTEIN, HETERODIMER 1FOS 19	
371	1fos	E	115	146	3.4e-10	-0.39	0.76		COMPLEX (GENE-REGULATORY PROTEIN/DNA) C-JUN PROTO-ONCOGENE (TRANSCRIPTION FACTOR AP-1) DIMERIZED 1FOS 4 WITH C-FOS AND COMPLEXED WITH DNA 1FOS 5 COILED-COIL, DNA-BINDING PROTEIN, HETERODIMER 1FOS 19	
373	1d5t	A	166	598	0	0.32	1.00		GUANINE NUCLEOTIDE DISSOCIATION INHIBITOR; CHAIN: A;	HYDROLASE INHIBITOR ULTRA-HIGH RESOLUTION
373	1qo8	A	8	46	0.0045	0.01	0.17		FLAVOCYTOCHROME C3 FUMARATE REDUCTASE; CHAIN: A, D;	OXIDOREDUCTASE OXIDOREDUCTASE
373	3lad	A	8	48	0.006	-0.12	0.36		OXIDOREDUCTASE DIHYDROLIPOAMIDE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									DEHYDROGENASE (E.C.1.8.1.4) 3LAD 3'	
374	1aih	A	168	252	5.1e-15	0.00	0.05		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;.	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	188	280	6.8e-22	-0.03	0.30		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	228	304	3.4e-23	0.60	0.12		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	308	388	1.2e-29	-0.01	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	308	389	1.2e-32	-0.32	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	336	416	1e-30	0.03	0.92		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	393	472	1.2e-37	0.64	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	420	502	1.2e-37			86.81	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	476	556	1.2e-34	0.57	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
374	1aih	A	476	556	1.7e-31	0.43	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									SITE; CHAIN: B, C;	
374	1mey	C	186	280	3.4e-38	0.45	0.75		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	227	301	8.5e-41	0.40	0.84		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	255	360	1e-43	-0.15	0.35		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	307	388	1e-48	0.06	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	335	416	5.1e-50	-0.05	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	363	444	1e-50	0.39	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	391	472	1.7e-51	0.48	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	419	500	6.8e-51	0.55	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	447	528	1.2e-50	0.51	1.00		DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	447	529	6.8e-51			106.37	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	C	475	556	1.7e-50	0.37	1.00		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1mey	G	225	252	1.5e-10	-0.12	0.69		DNA: CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
374	1tf3	A	187	276	6.8e-14	0.06	-0.06		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
374	1tf6	A	187	341	5.1e-29	0.05	-0.07		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1tf6	A	307	470	8.5e-39			117.85	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1tf6	A	308	453	6.8e-38	0.01	0.98		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1tf6	A	336	481	1.7e-38	0.12	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1tf6	A	392	538	8.5e-39	0.13	0.96		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1tf6	A	448	556	3.4e-30	0.18	0.46		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
374	1ubd	C	166	280	8.5e-25	0.10	0.05		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	190	304	3.4e-27	0.28	0.60		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	263	360	8.5e-29	-0.15	0.19		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										REGULATION/DNA)
374	1ubd	C	287	388	5.1e-34	0.12	0.94		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	312	444	9e-41	0.13	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	315	416	1.5e-34	0.01	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	343	444	1.5e-34	0.30	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	399	500	5.1e-36	0.23	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	418	529	1.5e-51	0.18	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	421	529	1.5e-51			98.87	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	445	556	1.5e-46	0.20	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1ubd	C	455	556	1.5e-34	0.21	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
374	1zfd		532	558	6.8e-05	0.06	0.30		SWI5; CHAIN: NULL;	ZINC FINGER DNA BINDING DOMAIN DNA BINDING MOTIF, ZINC FINGER DNA BINDING DOMAIN
374	2adr		189	254	3.4e-11	-0.04	0.06		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
374	2gli	A	161	303	8.5e-24	0.07	-0.11		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
374	2gli	A	287	415	1.2e-34	0.18	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										BINDING PROTEIN/DNA)
374	2gli	A	335	474	1.2e-61			106.08	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
374	2gli	A	393	530	1.2e-61	0.49	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
374	2gli	A	420	557	4.5e-58	0.36	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
375	1c3t	A	1	76	1e-31	0.68	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
375	1c3t	A	1	76	1e-31			102.61	1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
375	1tbe	B	1	72	1.2e-32	0.97	1.00		UBIQUITIN TETRAUBIQUITIN 1TBE 3	
375	1tbe	B	1	72	1.2e-32			97.63	UBIQUITIN TETRAUBIQUITIN 1TBE 3	
375	1ubi		1	76	1e-33	1.07	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
375	1ubi		1	76	7.5e-36			105.89	CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
375	1ubi		1	76	7.5e-36	1.07	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
375	1ud7	A	1	76	1.2e-32	0.96	1.00		UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
375	1ud7	A	1	76	1.2e-32			102.60	UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
377	1cdm	A	5	144	1.2e-62	0.90	1.00		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
377	1cdm	A	5	144	1.2e-62			149.72	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
377	1cll		5	144	3.4e-66	1.07	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
377	1cll		5	145	3.4e-66			156.05	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
377	1cmf		74	146	1.5e-23			79.20	CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
377	1cmf		81	143	1.5e-23	0.90	1.00		CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
377	1exr	A	3	143	5.1e-64	0.96	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
377	1f71	A	81	143	1.5e-23	1.14	1.00		CALMODULIN; CHAIN: A;	TRANSPORT PROTEIN CALCIUM BINDING, EF HAND, FOUR-HELIX BUNDLE
377	1tnx		1	143	3.4e-50			127.27	TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
377	1tnx		5	143	3.4e-50	0.85	1.00		TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
377	1vrk	A	2	146	1.5e-66	1.08	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
377	1vrk	A	2	146	1.5e-66			156.22	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
384	1b7f	A	2	113	1.7e-21	0.43	0.99		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									*UP*UP*UP*U)- CHAIN: P, Q;	
384	1b7f	A	33	205	3.4e-43	1.07	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
384	1b7f	A	33	205	3.4e-43			84.87	SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
384	1cvj	A	2	119	1.5e-31	0.42	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1cvj	A	37	211	1.4e-43	0.72	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1cvj	A	378	500	3.4e-23	0.16	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1cvj	B	2	99	6.8e-26	0.31	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1cvj	B	37	188	1.7e-37	0.57	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1cvj	F	37	178	8.5e-28	0.33	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	
384	1cvj	H	37	181	1.4e-28	0.46	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
384	1d8z	A	32	117	5.1e-22	0.61	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1d8z	A	419	501	4.5e-24	0.83	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1d9a	A	36	120	1.5e-17	0.77	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1d9a	A	418	501	4.5e-23	0.72	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1hal		30	205	1.7e-51	0.70	1.00		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
384	1hal		31	204	1.7e-51			74.92	HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
384	1hal		376	494	1e-23	0.63	-0.05		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
384	1hal		4	113	6.8e-22	0.33	0.63		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
384	1hal		413	498	3.4e-28	0.70	1.00		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
384	1hd1	A	36	113	1e-22	0.91	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1hd1	A	419	494	8.5e-24	1.02	0.99		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	1sxl		406	501	6e-25	0.48	0.99		RNA-BINDING PROTEIN SEX-LETHAL PROTEIN (C-TERMINUS, OR SECOND RNA-BINDING DOMAIN 1SXL 3 (RBD-2), RESIDUES 199 - 294 PLUS N-TERMINAL MET) 1SXL 4 (NMR, 17 STRUCTURES) 1SXL 5	
384	2mss	A	36	113	6.8e-18	0.50	0.58		MUSASHI1; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
384	2sxl		33	118	3.4e-20	0.63	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
384	2up1	A	29	210	1.4e-53	0.69	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
384	2up1	A	30	213	1.4e-53			77.86	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
384	2up1	A	376	499	1e-24	-0.07	0.06		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
384	2up1	A	4	119	5.1e-23	0.44	0.63		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
384	2upl	A	412	501	1.5e-29	0.87	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
384	3sxl	A	2	106	1.2e-20	0.47	0.99		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
384	3sxl	A	35	189	3.4e-41	0.72	1.00		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
391	1a06		1	327	1.7e-63			98.83	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
391	1a6o		1	296	1.2e-81			153.21	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
391	1a6o		3	295	1.2e-81	0.30	1.00		PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
391	1apm	E	1	324	6e-55			116.50	TRANSFERASE(PHOSPHOTRANSFERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$)	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
391	1apm	E	2	288	1e-53	0.45	1.00		TRANSFERASE(PHOSPHOTRANSFERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
391	1apm	E	2	304	6e-55	0.31	1.00		TRANSFERASE(PHOSPHOTRANSFERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
391	1aq1		2	294	0	0.37	1.00		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
391	1aq1		2	298	0			212.68	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
391	1bi8	A	3	289	3.4e-91			182.71	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
391	1bi8	A	4	289	3.4e-91	0.04	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: B, D:	DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
391	1blx	A	1	296	1.7e-99			202.88	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
391	1blx	A	4	291	1.7e-99	0.27	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
391	1byg	A	1	303	3e-34			74.19	C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2 STAUROSPORINE, TRANSFERASE
391	1cki	A	2	281	3e-55			68.61	CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
391	1cki	A	4	288	3e-55	0.17	0.89		CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
391	1cm8	A	1	326	0	0.42	1.00		PHOSPHORYLATED MAP KINASE P38-GAMMA: CHAIN: A, B:	TRANSFERASE STRESS-ACTIVATED PROTEIN KINASE-3, ERK6, ERK5; P38-GAMMA, GAMMA, PHOSPHORYLATION, MAP KINASE
391	1cmk	E	1	324	6.8e-56			111.92	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
391	1cmk	E	2	288	6.8e-56	0.46	1.00		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
391	1csn		1	284	5.1e-18			77.16	CASEIN KINASE-1; 1CSN 4	PHOSPHOTRANSFERASE
391	1ctp	E	1	311	1.5e-56			109.28	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
391	1f3m	C	3	297	7.5e-67	0.41	1.00		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
391	1fgk	A	1	299	1.5e-38			95.41	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
391	1fgk	B	1	298	7.5e-37			101.29	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
391	1hcl		2	294	0	0.67	1.00		HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
391	1hcl		2	298	0			239.66	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
391	1ian		1	328	0	0.12	1.00		P38 MAP KINASE; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE CSBP, RK, P38; PROTEIN SER/THR-KINASE, SERINE/THREONINE-PROTEIN KINASE
391	1ian		1	328	0			163.36	P38 MAP KINASE; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE CSBP, RK, P38; PROTEIN SER/THR-KINASE, SERINE/THREONINE-PROTEIN KINASE
391	1ir3	A	1	275	4.5e-37			79.01	INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
391	1jnk		1	323	0	0.46	1.00		C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
391	1jnk		1	331	0			161.78	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
391	1koa		1	302	1e-57	0.26	1.00		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
391	1koa		1	358	1e-57			86.80	TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
391	1kob	A	1	292	1.7e-57	0.26	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
391	1kob	A	1	357	1.7e-57			124.22	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
391	1p38		1	328	0	0.47	1.00		MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
391	1p38		1	332	0			191.19	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
391	1phk		1	291	1.7e-66			123.81	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
391	1phk		3	291	1.7e-66	0.37	1.00		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
391	1pmc		1	330	0	0.53	1.00		ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
391	1pme		1	331	0			183.19	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
391	1tki	A	1	358	1.7e-45			114.84	TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
391	3erk		1	325	0			187.32	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
391	3erk		1	326	0	0.54	1.00		EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
393	1apq		120	154	1.5e-11	-0.02	1.00		COMPLEMENT PROTEASE C1R: CHAIN: NULL;	COMPLEMENT COMPLEMENT, EGF, CALCIUM BINDING, SERINE PROTEASE
393	1ck4	A	5	111	6e-25	0.51	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
393	1ck4	A	527	709	1e-46	1.12	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
393	1dan	L	116	205	4.5e-20	-0.44	0.65		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	124	246	3e-32	-0.30	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	168	287	4.5e-31	-0.15	0.55		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	207	328	6e-31	-0.25	0.57		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	248	369	3e-25	-0.40	0.11		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	276	358	6.8e-16	-0.17	0.84		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	317	397	3.4e-16	-0.32	0.47		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	332	451	9e-25	-0.23	0.17		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	336	447	1.7e-18	-0.42	0.00		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	372	492	9e-26	-0.12	0.05		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	PROTEASE/COFACTOR/LIGAND)
393	1dan	L	412	535	1.2e-30	0.20	0.22		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	439	528	1.7e-17	-0.09	0.94		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dqb	A	438	525	7.5e-17	0.40	0.98		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
393	1dva	L	317	397	3.4e-16	-0.62	0.58		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
393	1dva	L	439	528	1.7e-17	0.21	0.84		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
393	1dx5	I	121	233	1e-23	0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	153	274	3e-25	0.09	0.55		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F,	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									G, H;	ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	195	315	4.5e-27	0.30	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	235	346	1.2e-17	0.02	0.99		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	236	356	1.5e-26	-0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	318	438	1.5e-22	0.19	0.93		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	359	479	3e-24	0.41	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	401	520	3e-24	0.61	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
393	1dx5	I	79	188	6.8e-15	-0.40	0.05		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1emn		273	347	5.1e-19	0.06	0.78		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		317	388	3.4e-18	-0.20	0.99		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		440	511	5.1e-18	0.32	1.00		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1fak	L	150	246	7.5e-22	-0.23	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	192	287	1.5e-21	-0.05	0.24		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	232	328	3e-23	0.08	0.80		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	273	369	3e-18	-0.27	0.15		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	276	358	6.8e-16	0.01	0.90		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	317	397	3.4e-16	-0.41	0.35		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	355	451	6e-19	0.35	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									FACTOR; CHAIN: T; 5L15; CHAIN: I;	PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	396	492	4.5e-19	0.09	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	437	527	3e-21	0.44	0.76		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	439	528	1.7e-17	0.22	0.98		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1ido		1	109	4.5e-24	0.33	0.84		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1ido		527	707	4.5e-46			98.35	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1ido		529	706	4.5e-46	1.04	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1jia	A	205	321	1.5e-19	0.01	-0.02		PHOSPHOLIPASE A2; CHAIN: A, B;	PHOSPHOLIPASE PHOSPHOLIPASE A2, AGKISTRODON HALYS PALLAS CRYSTAL 2 STRUCTURE
393	1lfa	A	1	112	1.5e-24	0.01	0.89		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1lfa	A	526	711	1.5e-53			93.64	CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1lfa	A	526	713	1.5e-53	1.12	1.00		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1pfx	L	157	301	4.5e-30	-0.01	0.07		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	197	341	3e-29	-0.15	0.81		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	286	423	3e-23	-0.10	0.06		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	403	527	6e-25	0.06	0.41		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	440	538	1.2e-15	-0.13	0.11		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1qfk	L	444	528	1.7e-16	0.30	0.86		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
393	1xka	L	444	528	1.7e-14	0.33	0.64		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
393	1apq		120	154	1.5e-11	-0.02	1.00		COMPLEMENT PROTEASE C1R; CHAIN: NULL;	COMPLEMENT COMPLEMENT, EGF, CALCIUM BINDING, SERINE PROTEASE
393	1aut	L	80	151	1.2e-10	-0.62	0.05		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
393	1ck4	A	5	111	6e-25	0.51	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
393	1ck4	A	527	709	1e-46	1.12	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
393	1dan	L	116	205	4.5e-20	-0.44	0.65		BLOOD COAGULATION FACTOR VIIA: CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	124	246	3e-32	-0.30	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	168	287	4.5e-31	-0.15	0.55		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	207	328	6e-31	-0.25	0.57		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	248	369	3e-25	-0.40	0.11		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	273	365	1.2e-16	0.02	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	332	451	9e-25	-0.23	0.17		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	372	492	9e-26	-0.12	0.05		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	PROTEASE/COFACTOR/LIGAND)
393	1dan	L	412	535	1.2e-30	0.20	0.22		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dan	L	439	528	1.7e-18	0.18	0.89		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
393	1dqb	A	315	401	1.1e-15	0.05	0.69		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
393	1dqb	A	438	525	7.5e-17	0.40	0.98		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
393	1dva	L	273	365	1.2e-16	-0.09	0.63		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
393	1dva	L	439	528	1.7e-18	0.32	0.92		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
393	1dx5	I	121	233	1e-23	0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	153	274	3e-25	0.09	0.55		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	195	315	4.5e-27	0.30	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	236	356	1.5e-26	-0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	316	438	3.4e-17	-0.08	0.99		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	318	438	1.5e-22	0.19	0.93		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	359	479	3e-24	0.41	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	401	520	3e-24	0.61	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	442	525	1.5e-13	0.45	0.78		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1dx5	I	77	188	3.4e-15	-0.52	0.18		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
393	1emn		112	187	3.4e-16	0.12	0.96		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		235	306	1.7e-18	0.27	0.81		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		273	347	1.7e-17	0.10	0.88		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		317	392	1e-17	-0.34	0.80		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1emn		710	779	6.8e-15	0.06	-0.19		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
393	1fak	L	107	164	6e-11	-0.12	0.31		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	150	246	7.5e-22	-0.23	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	192	287	1.5e-21	-0.05	0.24		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	232	328	3e-23	0.08	0.80		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	273	365	1.2e-16	-0.05	0.21		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	273	369	3e-18	-0.27	0.15		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	355	451	6e-19	0.35	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	396	492	4.5e-19	0.09	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	437	527	3e-21	0.44	0.76		BLOOD COAGULATION FACTOR VIIA: CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									FACTOR; CHAIN: T; 5L15; CHAIN: I;	PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1fak	L	439	528	1.7e-18	0.13	0.94		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
393	1ido		1	109	4.5e-24	0.33	0.84		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1ido		527	707	4.5e-46			98.35	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1ido		529	706	4.5e-46	1.04	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
393	1jia	A	205	321	1.5e-19	0.01	-0.02		PHOSPHOLIPASE A2; CHAIN: A, B;	PHOSPHOLIPASE PHOSPHOLIPASE A2, AGKISTRODON HALYS PALLAS CRYSTAL 2 STRUCTURE
393	1lfa	A	1	112	1.5e-24	0.01	0.89		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1lfa	A	526	711	1.5e-53			93.74	CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1lfa	A	526	713	1.5e-53	1.12	1.00		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
393	1pfx	L	157	301	4.5e-30	-0.01	0.07		FACTOR IXA; CHAIN: C, L,; D-	COMPLEX (BLOOD

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PHE-PRO-ARG; CHAIN: I;	COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	197	341	3e-29	-0.15	0.81		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	286	423	3e-23	-0.10	0.06		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	403	527	6e-25	0.06	0.41		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1pfx	L	440	536	8.5e-15	0.30	0.94		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
393	1qfk	L	444	528	3.4e-17	0.06	0.92		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									TRIPEPTIDYL INHIBITOR; CHAIN: C;	
393	1xka	L	444	528	1.5e-14	0.33	0.64		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
394	1apq		120	154	1.5e-11	-0.02	1.00		COMPLEMENT PROTEASE C1R; CHAIN: NULL;	COMPLEMENT COMPLEMENT, EGF, CALCIUM BINDING, SERINE PROTEASE
394	1ck4	A	5	111	6e-25	0.51	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
394	1ck4	A	527	709	1e-46	1.12	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
394	1dan	L	116	205	4.5e-20	-0.44	0.65		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	124	246	3e-32	-0.30	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	168	287	4.5e-31	-0.15	0.55		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	207	328	6e-31	-0.25	0.57		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	248	369	3e-25	-0.40	0.11		BLOOD COAGULATION FACTOR	BLOOD COAGULATION, SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	276	358	6.8e-16	-0.17	0.84		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	317	397	3.4e-16	-0.32	0.47		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	332	451	9e-25	-0.23	0.17		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	336	447	1.7e-18	-0.42	0.00		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	372	492	9e-26	-0.12	0.05		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	412	535	1.2e-30	0.20	0.22		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	439	528	1.7e-17	-0.09	0.94		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dqb	A	438	525	7.5e-17	0.40	0.98		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
394	1dva	L	317	397	3.4e-16	-0.62	0.58		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
394	1dva	L	439	528	1.7e-17	0.21	0.84		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
394	1dx5	I	121	233	1e-23	0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	153	274	3e-25	0.09	0.55		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	195	315	4.5e-27	0.30	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	235	346	1.2e-17	0.02	0.99		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	236	356	1.5e-26	-0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	318	438	1.5e-22	0.19	0.93		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	359	479	3e-24	0.41	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	401	520	3e-24	0.61	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	79	188	6.8e-15	-0.40	0.05		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1emn		273	347	5.1e-19	0.06	0.78		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		317	388	3.4e-18	-0.20	0.99		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		440	511	5.1e-18	0.32	1.00		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1fak	L	150	246	7.5e-22	-0.23	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	192	287	1.5e-21	-0.05	0.24		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	232	328	3e-23	0.08	0.80		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	273	369	3e-18	-0.27	0.15		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	276	358	6.8e-16	0.01	0.90		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	317	397	3.4e-16	-0.41	0.35		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	355	451	6e-19	0.35	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	396	492	4.5e-19	0.09	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN:	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									I;	RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	437	527	3e-21	0.44	0.76		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	439	528	1.7e-17	0.22	0.98		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1ido		1	109	4.5e-24	0.33	0.84		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1ido		527	707	4.5e-46			98.35	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1ido		529	706	4.5e-46	1.04	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1jia	A	205	321	1.5e-19	0.01	-0.02		PHOSPHOLIPASE A2; CHAIN: A, B;	PHOSPHOLIPASE PHOSPHOLIPASE A2, AGKISTRODON HALYS PALLAS CRYSTAL 2 STRUCTURE
394	1lfa	A	1	112	1.5e-24	0.01	0.89		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
394	1lfa	A	526	711	1.5e-53			93.64	CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
394	1lfa	A	526	713	1.5e-53	1.12	1.00		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
394	1pfx	L	157	301	4.5e-30	-0.01	0.07		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	197	341	3e-29	-0.15	0.81		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	286	423	3e-23	-0.10	0.06		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	403	527	6e-25	0.06	0.41		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	440	538	1.2e-15	-0.13	0.11		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1qfk	L	444	528	1.7e-16	0.30	0.86		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
394	1xka	L	444	528	1.7e-14	0.33	0.64		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
394	1apq		120	154	1.5e-11	-0.02	1.00		COMPLEMENT PROTEASE C1R; CHAIN: NULL;	COMPLEMENT COMPLEMENT, EGF, CALCIUM BINDING, SERINE PROTEASE
394	1aut	L	80	151	1.2e-10	-0.62	0.05		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
394	1ck4	A	5	111	6e-25	0.51	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
394	1ck4	A	527	709	1e-46	1.12	1.00		INTEGRIN ALPHA-1; CHAIN: A, B;	STRUCTURAL PROTEIN I-DOMAIN, METAL BINDING, COLLAGEN, ADHESION
394	1dan	L	116	205	4.5e-20	-0.44	0.65		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	124	246	3e-32	-0.30	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
394	1dan	L	168	287	4.5e-31	-0.15	0.55		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	207	328	6e-31	-0.25	0.57		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	248	369	3e-25	-0.40	0.11		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	273	365	1.2e-16	0.02	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	332	451	9e-25	-0.23	0.17		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	372	492	9e-26	-0.12	0.05		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	412	535	1.2e-30	0.20	0.22		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dan	L	439	528	1.7e-18	0.18	0.89		BLOOD COAGULATION FACTOR	BLOOD COAGULATION, SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
394	1dqb	A	315	401	1.1e-15	0.05	0.69		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
394	1dqb	A	438	525	7.5e-17	0.40	0.98		THROMBOMODULIN; CHAIN: A;	MEMBRANE PROTEIN NMR, THROMBIN, EGF MODULE, ANTICOAGULANT, GLYCOSYLATION
394	1dva	L	273	365	1.2e-16	-0.09	0.63		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
394	1dva	L	439	528	1.7e-18	0.32	0.92		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
394	1dx5	I	121	233	1e-23	0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	153	274	3e-25	0.09	0.55		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	195	315	4.5e-27	0.30	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J,	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	236	356	1.5e-26	-0.04	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	316	438	3.4e-17	-0.08	0.99		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	318	438	1.5e-22	0.19	0.93		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	359	479	3e-24	0.41	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	401	520	3e-24	0.61	1.00		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	442	525	1.5e-13	0.45	0.78		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1dx5	I	77	188	3.4e-15	-0.52	0.18		THROMBIN LIGHT CHAIN; CHAIN: A, B, C, D; THROMBIN HEAVY CHAIN; CHAIN: M, N, O, P; THROMBOMODULIN; CHAIN: I, J, K, L; THROMBIN INHIBITOR L-GLU-L-GLY-L-ARM; CHAIN: E, F, G, H;	SERINE PROTEINASE COAGULATION FACTOR II; COAGULATION FACTOR II; FETOMODULIN, TM, CD141 ANTIGEN; EGR-CMK SERINE PROTEINASE, EGF-LIKE DOMAINS, ANTICOAGULANT COMPLEX, 2 ANTIFIBRINOLYTIC COMPLEX
394	1emn		112	187	3.4e-16	0.12	0.96		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		235	306	1.7e-18	0.27	0.81		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		273	347	1.7e-17	0.10	0.88		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		317	392	1e-17	-0.34	0.80		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
394	1emn		710	779	6.8e-15	0.06	-0.19		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MATRIX PROTEIN
394	1fak	L	107	164	6e-11	-0.12	0.31		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	150	246	7.5e-22	-0.23	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	192	287	1.5e-21	-0.05	0.24		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	232	328	3e-23	0.08	0.80		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	273	365	1.2e-16	-0.05	0.21		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	273	369	3e-18	-0.27	0.15		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	355	451	6e-19	0.35	0.23		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA, CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	396	492	4.5e-19	0.09	0.10		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	437	527	3e-21	0.44	0.76		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1fak	L	439	528	1.7e-18	0.13	0.94		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									FACTOR; CHAIN: T; 5L15; CHAIN: I;	PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
394	1ido		1	109	4.5e-24	0.33	0.84		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1ido		527	707	4.5e-46			98.35	INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1ido		529	706	4.5e-46	1.04	1.00		INTEGRIN; CHAIN: NULL;	CELL ADHESION PROTEIN A-DOMAIN INTEGRIN, CELL ADHESION PROTEIN, GLYCOPROTEIN, EXTRACELLULAR 2 MATRIX, CYTOSKELETON
394	1jia	A	205	321	1.5e-19	0.01	-0.02		PHOSPHOLIPASE A2; CHAIN: A, B;	PHOSPHOLIPASE PHOSPHOLIPASE A2, AGKISTRODON HALYS PALLAS CRYSTAL 2 STRUCTURE
394	1lfa	A	1	112	1.5e-24	0.01	0.89		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
394	1lfa	A	526	711	1.5e-53			93.74	CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
394	1lfa	A	526	713	1.5e-53	1.12	1.00		CD11A; 1LFA 5 CHAIN: A, B; 1LFA 6	CELL ADHESION LFA-1, ALPHA-L\BETA-2 INTEGRIN, A-DOMAIN; 1LFA 8
394	1pfx	L	157	301	4.5e-30	-0.01	0.07		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	197	341	3e-29	-0.15	0.81		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	286	423	3e-23	-0.10	0.06		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	403	527	6e-25	0.06	0.41		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1pfx	L	440	536	8.5e-15	0.30	0.94		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
394	1qfk	L	444	528	3.4e-17	0.06	0.92		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
394	1xka	L	444	528	1.5e-14	0.33	0.64		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
399	1aip	A	183	403	3.4e-67	-0.15	0.17		ELONGATION FACTOR TU;	COMPLEX OF TWO ELONGATION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, B, E, F; ELONGATION FACTOR TS; CHAIN: C, D, G, H;	FACTORS EF-TU; EF-TS; ELONGATION FACTOR, NUCLEOTIDE EXCHANGE, GTP-BINDING, 2 COMPLEX OF TWO ELONGATION FACTORS
399	1cfc	A	183	403	3.4e-71	-0.23	0.01		ELONGATION FACTOR; CHAIN: A, B;	RNA BINDING PROTEIN EFTU; TRANSPORT AND PROTECTION PROTEIN, RNA BINDING PROTEIN
399	1efu	A	183	403	5.1e-65	-0.21	0.09		ELONGATION FACTOR TU; CHAIN: A, C; ELONGATION FACTOR TS; CHAIN: B, D;	COMPLEX (TWO ELONGATION FACTORS) ELONGATION FACTOR FOR TRANSFER, HEAT UNSTABLE, ELONGATION FACTOR FOR TRANSFER, HEAT STABLE, ELONGATION FACTOR, COMPLEX (TWO ELONGATION FACTORS)
399	1ega	A	184	388	6.8e-38	-0.11	0.13		GTP-BINDING PROTEIN ERA; CHAIN: A, B;	HYDROLASE ERA, GTPASE, RNA-BINDING, RAS-LIKE, HYDROLASE
399	1etu		183	343	5.1e-47	0.07	0.41		TRANSPORT AND PROTECTION PROTEIN ELONGATION FACTOR TU (DOMAIN I) - *GUANOSINE DIPHOSPHATE 1ETU 4 COMPLEX 1ETU 5	
399	1exm	A	183	403	1.7e-73	-0.28	0.10		ELONGATION FACTOR TU (EF-TU); CHAIN: A;	TRANSLATION EF-TU; GTPASE, MOLECULAR SWITCH, TRNA, RIBOSOME, Q-BETA REPLICASE, 2 CHAPERONE, DISULFIDE ISOMERASE
399	1f60	A	183	400	1e-73	-0.34	0.07		ELONGATION FACTOR EEF1A; CHAIN: A; ELONGATION FACTOR EEF1BA; CHAIN: B;	TRANSLATION PROTEIN-PROTEIN COMPLEX
399	1kao		184	342	1.7e-05	-0.06	0.13		RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
402	1a1h	A	167	239	1.5e-20	-0.09	0.27		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
402	1a1h	A	186	271	1.5e-20			58.92	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
402	1a1h	A	243	310	3.4e-24	0.08	1.00		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
402	1mey	C	166	239	5.1e-37	-0.10	0.30		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
402	1mey	C	185	269	1.4e-44	-0.23	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
402	1mey	C	214	300	1.4e-44			67.85	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
402	1mey	C	242	310	1e-37	-0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
402	1spl		273	301	0.00015	-0.17	0.99		SP1F3: CHAIN: NULL;	ZINC FINGER TRANSCRIPTION FACTOR SP1; ZINC FINGER, TRANSCRIPTION ACTIVATION, SP1
402	1tf3	A	214	303	3.4e-20			67.04	TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
402	1tf3	A	243	307	3.4e-20	0.06	0.48		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
402	1tf6	A	108	295	5.1e-39			78.75	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
402	1tf6	A	167	306	5.1e-39	-0.11	0.39		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
402	1ubd	C	167	269	1.7e-33	-0.23	0.31		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
402	1ubd	C	187	300	5.1e-49			165.14	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
402	1ubd	C	190	299	5.1e-49	0.13	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
402	1ubd	C	222	310	1.7e-30	-0.28	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										REGULATION/DNA)
402	2gli	A	149	301	8.5e-38			82.64	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
402	2gli	A	167	298	8.5e-38	-0.15	0.94		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
404	1fjg	Q	71	147	1.5e-28	0.12	0.99		16S RIBOSOMAL RNA; CHAIN: A; FRAGMENT OF MESSENGER RNA; CHAIN: X; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL PROTEIN S8; CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14; CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T; 30S RIBOSOMAL PROTEIN THX; CHAIN: V	RIBOSOME 30S RIBOSOMAL SUBUNIT, RIBOSOME, ANTIBIOTIC, STREPTOMYCIN, 2 SPECTINOMYCIN, PAROMOMYCIN
404	1qd7	I	69	151	3.4e-32	-0.76	0.00		CENTRAL FRAGMENT OF 16 S RNA; CHAIN: A; END FRAGMENT	RIBOSOME 30S RIBOSOMAL SUBUNIT, LOW RESOLUTION MODEL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									OF 16 S RNA; CHAIN: B; S4 RIBOSOMAL PROTEIN; CHAIN: C; S5 RIBOSOMAL PROTEIN; CHAIN: D; S6 RIBOSOMAL PROTEIN; CHAIN: E; S7 RIBOSOMAL PROTEIN; CHAIN: F; S8 RIBOSOMAL PROTEIN; CHAIN: G; S15 RIBOSOMAL PROTEIN; CHAIN: H; S17 RIBOSOMAL PROTEIN; CHAIN: I; S20 RIBOSOMAL PROTEIN; CHAIN: J	
406	1aps		2	98	1.4e-33	0.96	1.00		HYDROLASE(ACTING ON ACID ANHYDRIDES) ACYLPHOSPHATASE (E.C.3.6.1.7) (NMR, 5 STRUCTURES) 1APS 3	
406	1aps		2	99	1.4e-33			102.47	HYDROLASE(ACTING ON ACID ANHYDRIDES) ACYLPHOSPHATASE (E.C.3.6.1.7) (NMR, 5 STRUCTURES) 1APS 3	
406	2acy		2	99	3.4e-33	0.79	1.00		ACYLPHOSPHATASE; CHAIN: NULL;	ACYLPHOSPHATASE ACP; ACYLPHOSPHATASE, PHOSPHORIC MONOESTER HYDROLASE
406	2acy		2	99	3.4e-33			139.55	ACYLPHOSPHATASE; CHAIN: NULL;	ACYLPHOSPHATASE ACP; ACYLPHOSPHATASE, PHOSPHORIC MONOESTER HYDROLASE
407	1a17		622	730	1.5e-11	0.15	0.77		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
407	1a17		661	728	5.1e-06	0.09	0.98		SERINE/THREONINE PROTEIN PHOSPHATASE 5; CHAIN: NULL;	HYDROLASE TETRATRICOPEPTIDE, TRP; HYDROLASE, PHOSPHATASE, PROTEIN-PROTEIN INTERACTIONS, TPR, 2 SUPER-HELIX, X-RAY STRUCTURE
407	1elr	A	263	376	1.2e-07	-0.04	0.12		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
407	1elr	A	620	727	9e-10	-0.47	0.00		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: B;	REPEAT, HSP90, 2 PROTEIN BINDING
407	1elr	A	660	733	6.8e-05	-0.38	0.40		TPR2A-DOMAIN OF HOP; CHAIN: A; HSP90-PEPTIDE MEEVD; CHAIN: B;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSP90, 2 PROTEIN BINDING
407	1elw	A	658	758	1.2e-07	-0.23	0.21		TPR1-DOMAIN OF HOP; CHAIN: A, B; HSC70-PEPTIDE; CHAIN: C, D;	CHAPERONE HOP, TPR-DOMAIN, PEPTIDE-COMPLEX, HELICAL REPEAT, HSC70, 2 HSP70, PROTEIN BINDING
407	1fch	A	192	386	6e-12	-0.22	0.24		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
407	1fch	A	510	749	5.1e-12	-0.31	0.09		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
407	1fch	A	550	840	5.1e-15	-0.44	0.07		PEROXISOMAL TARGETING SIGNAL 1 RECEPTOR; CHAIN: A, B; PTS1-CONTAINING PEPTIDE; CHAIN: C, D;	SIGNALING PROTEIN PEROXISMORE RECEPTOR 1, PTS1-BP, PEROXIN-5, PTS1 PROTEIN-PEPTIDE COMPLEX, TETRATRICOPEPTIDE REPEAT, TPR, 2 HELICAL REPEAT
407	4hb1		703	744	0.0036	-0.01	0.10		DHP1; CHAIN: NULL;	DESIGNED HELICAL BUNDLE DESIGNED HELICAL BUNDLE
414	1a5j		112	146	0.00075	-0.08	0.62		B-MYB; CHAIN: NULL;	DNA-BINDING PROTEIN DNA-BINDING PROTEIN, PROTOONCOGENE PRODUCT
414	1ak2		749	973	1.7e-52			304.37	ADENYLATE KINASE ISOENZYME-2; CHAIN: NULL;	PHOSPHOTRANSFERASE ATP\:AMP PHOSPHOTRANSFERASE, MYOKINASE; NUCLEOSIDE MONOPHOSPHATE KINASE, PHOSPHOTRANSFERASE
414	1ak2		756	972	1.7e-52	0.84	1.00		ADENYLATE KINASE ISOENZYME-2; CHAIN: NULL;	PHOSPHOTRANSFERASE ATP\:AMP PHOSPHOTRANSFERASE, MYOKINASE; NUCLEOSIDE MONOPHOSPHATE KINASE, PHOSPHOTRANSFERASE
414	1aky		751	971	4.5e-78			212.52	ADENYLATE KINASE; 1AKY 4 CHAIN: NULL; 1AKY 5	TRANSFERASE (PHOSPHOTRANSFERASE) ATP\:AMP PHOSPHOTRANSFERASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MYOKINASE; 1AKY 6 ATP:AMP PHOSPHOTRANSFERASE, MYOKINASE 1AKY 15
414	1aky		767	970	4.5e-78	0.66	1.00		ADENYLATE KINASE; 1AKY 4 CHAIN: NULL; 1AKY 5	TRANSFERASE (PHOSPHOTRANSFERASE) ATP\:AMP PHOSPHOTRANSFERASE, MYOKINASE; 1AKY 6 ATP:AMP PHOSPHOTRANSFERASE, MYOKINASE 1AKY 15
414	1e4v	A	767	967	1.5e-74	0.13	1.00		ADENYLATE KINASE; CHAIN: A;	TRANSFERASE(PHOSPHOTRANSFERASE) TRANSFERASE(PHOSPHOTRANSFERASE)
414	1mbj		113	146	7.5e-05	-0.18	0.51		MYB PROTO-ONCOGENE PROTEIN; 1MBJ 4	DNA BINDING PROTEIN PROTOONCOGENE PRODUCT 1MBJ 12
414	1mse	C	113	146	0.0015	-0.06	0.55		COMPLEX (BINDING PROTEIN/DNA) C-MYB DNA-BINDING DOMAIN COMPLEXED WITH DNA 1MSE 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1MSE 4 1MSE 84	
415	1e7u	A	3501	3986	1e-68	0.10	0.86		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K, PI 3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K, WORTMANNIN
415	1e8y	A	3501	3986	3.4e-68	0.02	1.00		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K
415	3fap	B	3581	3674	1.4e-24	0.05	-0.18		FK506-BINDING PROTEIN; CHAIN: A; FKBP12-RAPAMYCIN ASSOCIATED PROTEIN; CHAIN: B;	CELL CYCLE FKBP12; FRAP FKBP12, FRAP, RAPAMYCIN, COMPLEX, GENE THERAPY
415	1e7u	A	3480	4043	8.5e-83	-0.12	0.37		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K, PI 3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MESSENGER 2 GENERATION, PI3K, PI 3K, WORTMANNIN
415	1e8y	A	3480	4043	1e-77	0.13	0.94		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K
415	3fap	B	3581	3673	1.5e-21	0.07	-0.18		FK506-BINDING PROTEIN; CHAIN: A; FKBP12-RAPAMYCIN ASSOCIATED PROTEIN; CHAIN: B;	CELL CYCLE FKBP12; FRAP FKBP12, FRAP, RAPAMYCIN, COMPLEX, GENE THERAPY
416	1e7u	A	3501	3986	1e-68	0.10	0.86		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K, PI 3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K, WORTMANNIN
416	1e8y	A	3501	3986	3.4e-68	0.02	1.00		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K
416	3fap	B	3581	3674	1.4e-24	0.05	-0.18		FK506-BINDING PROTEIN; CHAIN: A; FKBP12-RAPAMYCIN ASSOCIATED PROTEIN; CHAIN: B;	CELL CYCLE FKBP12; FRAP FKBP12, FRAP, RAPAMYCIN, COMPLEX, GENE THERAPY
416	1e7u	A	3480	4043	8.5e-83	-0.12	0.37		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K, PI 3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K, WORTMANNIN
416	1e8y	A	3480	4043	1e-77	0.13	0.94		PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT; CHAIN: A;	PHOSPHOINOSITIDE 3-KINASE GAMMA PTDINS-3-KINASE P110, PI3K; PHOSPHOINOSITIDE 3-KINASE GAMMA, SECONDARY MESSENGER 2 GENERATION, PI3K, PI 3K
416	3fap	B	3581	3673	1.5e-21	0.07	-0.18		FK506-BINDING PROTEIN; CHAIN: A; FKBP12-RAPAMYCIN ASSOCIATED PROTEIN; CHAIN: B;	CELL CYCLE FKBP12; FRAP FKBP12, FRAP, RAPAMYCIN, COMPLEX, GENE THERAPY

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
418	1aip	A	181	384	1.7e-46	0.07	-0.15		ELONGATION FACTOR TU; CHAIN: A, B, E, F; ELONGATION FACTOR TS; CHAIN: C, D, G, H;	COMPLEX OF TWO ELONGATION FACTORS EF-TU; EF-TS; ELONGATION FACTOR, NUCLEOTIDE EXCHANGE, GTP-BINDING, 2 COMPLEX OF TWO ELONGATION FACTORS
418	1d2e	A	181	386	1.7e-44	0.32	-0.17		ELONGATION FACTOR TU (EF-TU); CHAIN: A, B, C, D	RNA BINDING PROTEIN G-PROTEIN, BETA-BARREL
418	1e0s	A	185	312	3e-05	0.05	0.07		ADP-RIBOSYLATION FACTOR 6; CHAIN: A;	G PROTEIN G PROTEIN, RAS, ARF, ARF6, MEMBRANE TRAFFIC
418	1efc	A	181	386	3.4e-50	0.20	-0.17		ELONGATION FACTOR; CHAIN: A, B;	RNA BINDING PROTEIN EFTU; TRANSPORT AND PROTECTION PROTEIN, RNA BINDING PROTEIN
418	1efu	A	181	386	5.1e-46	0.15	-0.17		ELONGATION FACTOR TU; CHAIN: A, C; ELONGATION FACTOR TS; CHAIN: B, D;	COMPLEX (TWO ELONGATION FACTORS) ELONGATION FACTOR FOR TRANSFER, HEAT UNSTABLE, ELONGATION FACTOR FOR TRANSFER, HEAT STABLE, ELONGATION FACTOR, COMPLEX (TWO ELONGATION FACTORS)
418	1ega	A	186	381	3.4e-36	0.10	0.01		GTP-BINDING PROTEIN ERA; CHAIN: A, B;	HYDROLASE ERA, GTPASE, RNA-BINDING, RAS-LIKE, HYDROLASE
418	1ega	A	34	185	8.5e-13	0.23	-0.19		GTP-BINDING PROTEIN ERA; CHAIN: A, B;	HYDROLASE ERA, GTPASE, RNA-BINDING, RAS-LIKE, HYDROLASE
418	1exm	A	179	384	5.1e-52	0.23	-0.17		ELONGATION FACTOR TU (EF-TU); CHAIN: A;	TRANSLATION EF-TU; GTPASE, MOLECULAR SWITCH, TRNA, RIBOSOME, Q-BETA REPLICASE, 2 CHAPERONE, DISULFIDE ISOMERASE
418	1f60	A	179	386	3.4e-31	0.23	-0.12		ELONGATION FACTOR EEF1A; CHAIN: A; ELONGATION FACTOR EEF1BA; CHAIN: B;	TRANSLATION PROTEIN-PROTEIN COMPLEX
418	1hur	A	185	312	9e-05	0.06	0.12		HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B: 1HUR 7	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
421	1a1h	A	201	281	1.2e-26	-0.10	0.99		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
421	1c2a	A	396	513	1e-09	0.14	-0.15		BOWMAN-BIRK TRYPSIN	HYDROLASE INHIBITOR ALL-BETA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									INHIBITOR; CHAIN: A	STRUCTURE, HYDROLASE INHIBITOR
421	1mey	C	172	253	6.8e-41	-0.21	0.58		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	200	281	6.8e-44	0.09	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	228	309	3.4e-46	0.64	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	256	337	1.4e-47	0.60	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	284	365	1.7e-48	0.55	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	312	393	3.4e-49	0.50	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	340	421	6.8e-49	0.61	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	368	449	5.1e-50	0.34	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
421	1mey	C	396	477	3.4e-51	0.56	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	424	505	5.1e-51	0.53	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN: CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	452	533	6.8e-51	0.42	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	452	534	5.1e-51			108.34	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	480	561	1.7e-50	0.40	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	508	589	8.5e-51	0.63	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	536	617	1.5e-50	0.31	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1mey	C	564	641	5.1e-46	0.13	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
421	1tf6	A	201	346	5.1e-35	0.01	0.96		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
421	1tf6	A	257	402	1.4e-36	0.24	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
421	1tf6	A	369	514	1.7e-38	0.25	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
421	1tf6	A	396	559	1.7e-38			118.07	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
421	1tf6	A	481	627	5.1e-38	0.04	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
421	1ubd	C	182	309	3e-26	-0.20	0.18		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	203	309	3.4e-31	0.06	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
421	1ubd	C	228	337	3e-51	0.33	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	282	393	4.5e-53	0.48	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	284	394	4.5e-53			92.65	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	320	421	1.2e-33	0.29	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	348	449	1.7e-34	0.15	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	366	477	3e-53	0.23	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, B;	INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	376	477	3.4e-36	0.36	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	401	505	8.5e-36	0.16	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	422	533	6e-56	0.28	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	450	562	3e-55	0.16	0.96		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	460	561	1.7e-35	0.13	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										(TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	478	589	3e-53	0.04	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	506	617	3e-53	0.20	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	516	617	5.1e-34	0.34	0.96		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	1ubd	C	534	641	1.1e-39	0.17	0.98		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
421	2gli	A	172	308	1.5e-31	-0.27	0.78		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	192	311	3e-41	0.06	0.98		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	228	367	1.4e-63	0.72	1.00		ZINC FINGER PROTEIN GLI1;	COMPLEX (DNA-BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A; DNA; CHAIN: C, D;	PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	264	392	1.7e-33	0.54	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	284	423	1.4e-63			110.65	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	312	535	1.5e-67	-0.12	0.75		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	320	448	3.4e-34	0.31	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	404	532	3.4e-34	0.47	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	452	591	1.5e-70	0.39	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
421	2gli	A	508	626	7.5e-55	0.24	0.94		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
423	1bcc	A	24	459	0	0.90	1.00		UBIQUINOL CYTOCHROME C OXIDOREDUCTASE; CHAIN: A, B, C, D, E, F, G, H, I, J;	OXIDOREDUCTASE CYTOCHROME BC1 COMPLEX, COMPLEX III; UBIQUINONE, OXIDOREDUCTASE, REDOX ENZYME, MEMBRANE PROTEIN, 2 RESPIRATORY CHAIN, ELECTRON TRANSPORT
423	1bcc	A	49	459	0			457.94	UBIQUINOL CYTOCHROME C OXIDOREDUCTASE; CHAIN: A, B, C, D, E, F, G, H, I, J;	OXIDOREDUCTASE CYTOCHROME BC1 COMPLEX, COMPLEX III; UBIQUINONE, OXIDOREDUCTASE, REDOX ENZYME, MEMBRANE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										PROTEIN, 2 RESPIRATORY CHAIN, ELECTRON TRANSPORT
423	1qcr	A	24	459	0	0.41	1.00		UBIQUINOL CYTOCHROME C OXIDOREDUCTASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K;	OXIDOREDUCTASE CYTOCHROME BC1, QCR; BC1, QCR, MEMBRANE PROTEIN, PROTON TRANSLOCATION, ELECTRON 2 TRANSFER, PROTEASE, MPP, MITOCHONDRIAL PROCESSING 3 PEPTIDASE, STRUCTURE, CYTOCHROME C1, CYTOCHROME B, RIESKE, 4 IRON SULFER PROTEIN, OXIDOREDUCTASE
426	1deq	B	122	285	1.4e-52	-0.25	0.62		FIBRINOGEN (ALPHA CHAIN); CHAIN: A, D, N, Q; FIBRINOGEN (BETA CHAIN); CHAIN: B, E, O, R; FIBRINOGEN (GAMMA CHAIN); CHAIN: C, F, P, S; FIBRINOGEN; CHAIN: M, Z;	BLOOD CLOTTING COILED-COIL
426	1deq	C	53	276	4.2e-89	-0.52	1.00		FIBRINOGEN (ALPHA CHAIN); CHAIN: A, D, N, Q; FIBRINOGEN (BETA CHAIN); CHAIN: B, E, O, R; FIBRINOGEN (GAMMA CHAIN); CHAIN: C, F, P, S; FIBRINOGEN; CHAIN: M, Z;	BLOOD CLOTTING COILED-COIL
426	1deq	C	53	286	8.5e-45	-0.58	1.00		FIBRINOGEN (ALPHA CHAIN); CHAIN: A, D, N, Q; FIBRINOGEN (BETA CHAIN); CHAIN: B, E, O, R; FIBRINOGEN (GAMMA CHAIN); CHAIN: C, F, P, S; FIBRINOGEN; CHAIN: M, Z;	BLOOD CLOTTING COILED-COIL
426	1ei3	C	29	286	3.4e-52	-0.58	1.00		FIBRINOGEN; CHAIN: A, D; FIBRINOGEN; CHAIN: B, E; FIBRINOGEN; CHAIN: C, F;	BLOOD CLOTTING COILED COILS, DISULFIDE RINGS, FIBRIN FORMING ENTITIES
426	1fzc	C	123	286	3.4e-39	0.19	1.00		FIBRIN; CHAIN: A, B, C, D, E, F, G, H, I, J;	BLOOD COAGULATION BLOOD COAGULATION, PLASMA PROTEIN, CROSSLINKING
426	1fzc	C	123	288	3.4e-39			175.96	FIBRIN; CHAIN: A, B, C, D, E, F, G, H, I, J;	BLOOD COAGULATION BLOOD COAGULATION, PLASMA PROTEIN, CROSSLINKING
426	1fzg	C	128	288	1e-38			174.90	FIBRINOGEN; CHAIN: A, B, C, D, E, F, S, T, M, N;	BLOOD COAGULATION BLOOD COAGULATION, PLASMA,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										PLATELET, FIBRINOGEN, FIBRIN
426	1fzg	C	129	286	1e-38	0.22	1.00		FIBRINOGEN; CHAIN: A, B, C, D, E, F, S, T, M, N;	BLOOD COAGULATION BLOOD COAGULATION, PLASMA, PLATELET, FIBRINOGEN, FIBRIN
432	2dnj	A	21	251	3.4e-100	0.93	1.00		ENDONUCLEASE DEOXYRIBONUCLEASE I (DNASE I) (E.C.3.1.21.1) COMPLEXED WITH 2DNJ 3 DNA (5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3') 2DNJ 4	
432	2dnj	A	21	252	3.4e-100			202.60	ENDONUCLEASE DEOXYRIBONUCLEASE I (DNASE I) (E.C.3.1.21.1) COMPLEXED WITH 2DNJ 3 DNA (5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3') 2DNJ 4	
433	1b50	A	25	92	1.1e-28			90.96	MIP-1A; CHAIN: A, B;	CHEMOKINE CHEMOKINE, CYTOKINE, CHEMOTAXIS
433	1b50	A	26	92	1.1e-28	0.01	1.00		MIP-1A; CHAIN: A, B;	CHEMOKINE CHEMOKINE, CYTOKINE, CHEMOTAXIS
433	1b50	A	27	92	5.1e-25	0.29	1.00		MIP-1A; CHAIN: A, B;	CHEMOKINE CHEMOKINE, CYTOKINE, CHEMOTAXIS
433	1hum	A	24	92	6.8e-25			114.82	CYTOKINE(CHEMOTACTIC) HUMAN MACROPHAGE INFLAMMATORY PROTEIN 1 BETA (HMIP-1B) 1HUM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1HUM 4	
433	1hum	A	25	92	6.8e-25	0.28	1.00		CYTOKINE(CHEMOTACTIC) HUMAN MACROPHAGE INFLAMMATORY PROTEIN 1 BETA (HMIP-1B) 1HUM 3 (NMR, MINIMIZED AVERAGE STRUCTURE) 1HUM 4	
433	1ncv	A	24	92	1.7e-25			61.57	MONOCYTE CHEMOATTRACTANT PROTEIN 3; CHAIN: A, B;	CYTOKINE NMR, STRUCTURE, MCP-3, BETA-CHEMOKINE, CYTOKINE, CHEMOTAXIS, 2 HEPARIN-BINDING, GLYCOPROTEIN
433	1ncv	A	25	91	1.7e-25	-0.04	0.98		MONOCYTE CHEMOATTRACTANT PROTEIN 3;	CYTOKINE NMR, STRUCTURE, MCP-3, BETA-CHEMOKINE, CYTOKINE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									CHAIN: A, B;	CHEMOTAXIS, 2 HEPARIN-BINDING, GLYCOPROTEIN
449	1awc	B	114	270	1.7e-39			61.04	GA BINDING PROTEIN ALPHA; CHAIN: A; GA BINDING PROTEIN BETA 1; CHAIN: B; DNA; CHAIN: D, E;	COMPLEX (TRANSCRIPTION REGULATION/DNA) GABPALPHA; GABPBETA1; COMPLEX (TRANSCRIPTION REGULATION/DNA), DNA-BINDING, 2 NUCLEAR PROTEIN, ETS DOMAIN, ANKYRIN REPEATS, TRANSCRIPTION 3 FACTOR
449	1bd8		116	273	8.4e-33			57.66	P19INK4D CDK4/6 INHIBITOR; CHAIN: NULL;	TUMOR SUPPRESSOR TUMOR SUPPRESSOR, CDK4/6 INHIBITOR, ANKYRIN MOTIF
449	1blx	B	115	276	1.4e-32			53.46	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
449	1bu9	A	113	280	3.4e-33			51.20	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A;	HORMONE/GROWTH FACTOR P18-INK4C; CELL CYCLE INHIBITOR, P18INK4C, TUMOR, SUPPRESSOR, CYCLIN- 2 DEPENDENT KINASE, HORMONE/GROWTH FACTOR
449	1ihb	A	122	273	1.5e-32			53.58	CYCLIN-DEPENDENT KINASE 6 INHIBITOR; CHAIN: A, B;	CELL CYCLE INHIBITOR P18-INK4C(INK6); CELL CYCLE INHIBITOR, P18-INK4C(INK6), ANKYRIN REPEAT, 2 CDK 4/6 INHIBITOR
449	1ikn	D	81	293	2.8e-44			62.44	NF-KAPPA-B P65 SUBUNIT; CHAIN: A; NF-KAPPA-B P50D SUBUNIT; CHAIN: C; I-KAPPA-B-ALPHA; CHAIN: D;	TRANSCRIPTION FACTOR P65; P50D; TRANSCRIPTION FACTOR, IKB/NFKB COMPLEX
449	1myo		48	166	9.8e-27			54.85	MYOTROPHIN; CHAIN: NULL	ANK-REPEAT MYOTROPHIN, ACETYLATION, NMR, ANK-REPEAT
449	1nfi	E	78	282	2.8e-44			63.92	NF-KAPPA-B P65; CHAIN: A, C; NF-KAPPA-B P50; CHAIN: B, D; I-KAPPA-B-ALPHA; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REG/ANK REPEAT) COMPLEX (TRANSCRIPTION REGULATION/ANK REPEAT), ANKYRIN 2 REPEAT HELIX
451	1ndh		36	305	5.1e-79			366.08	ELECTRON TRANSPORT (FLAVO	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
									PROTEIN) CYTOCHROME B=5= REDUCTASE (E.C.1.6.2.2) 1NDH 3	
456	1tub	A	2	440	0			308.77	TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
456	1tub	B	2	440	0			353.89	TUBULIN; CHAIN: A, B;	MICROTUBULES MICROTUBULES, ALPHA-TUBULIN, BETA-TUBULIN, GTPASE HELIX
457	1klo		82	239	2.8e-34			138.52	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
458	1bih	A	1	327	6.8e-47			77.74	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
458	1fig	H	54	276	0.00034			61.84	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (KAPPA LIGHT CHAIN) FAB' FRAGMENT 1FIG 3	
458	1for	H	64	278	0.0019			60.01	IMMUNOGLOBULIN IGG2A FAB FRAGMENT (FAB17-1A) (ORTHORHOMBIC CRYSTAL FORM) 1FOR 3	
458	1igc	H	58	279	0.00017			61.96	COMPLEX (ANTIBODY/BINDING PROTEIN) IGG1 FAB FRAGMENT COMPLEXED WITH PROTEIN G (DOMAIN III) 1IGC 5 PROTEIN G, STREPTOCOCCUS 1IGC 15	
458	1itb	B	1	279	4.2e-25			62.51	INTERLEUKIN-1 BETA; CHAIN: A; TYPE I INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
458	1kb5	H	54	278	0.0024			63.34	KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
458	2gfb	B	58	279	0.00034			67.09	IMMUNOGLOBULIN IGG2A FAB FRAGMENT (CNJ206) 2GFB 3	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
458	7fab	L	65	260	1.5e-11			58.17	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
462	1au7	A	143	289	3.4e-33			105.92	PIT-1; CHAIN: A, B; DNA; CHAIN: C, D:	COMPLEX (DNA-BINDING PROTEIN/DNA) GHF-1; COMPLEX (DNA-BINDING PROTEIN/DNA), PITUITARY, CPHD, 2 POU DOMAIN, TRANSCRIPTION FACTOR
462	1ocp		223	289	2.8e-22			84.91	OCT-3; 1OCP 5 CHAIN: NULL; 1OCP 6	DNA-BINDING PROTEIN
462	1oct	C	143	290	1.3e-40			120.80	DNA-BINDING PROTEIN OCT-1 (POU DOMAIN) 1OCT 3	
462	1pou		143	212	5.6e-32			79.90	DNA-BINDING PROTEIN OCT-1 (POU-SPECIFIC DOMAIN) (NMR, 20 STRUCTURES) 1POU 3	
473	1fht		30	143	2.8e-16			53.05	U1 SMALL NUCLEAR RIBONUCLEOPROTEIN A; CHAIN: NULL;	RIBONUCLEOPROTEIN U1A117; RIBONUCLEOPROTEIN, RNP DOMAIN, SPLICEOSOME
476	1c96	A	82	963	0			253.82	MITOCHONDRIAL ACONITASE; CHAIN: A;	LYASE CITRATE HYDRO-LYASE; LYASE, TRICARBOXYLIC ACID CYCLE, IRON-SULFUR, MITOCHONDRION, 2 TRANSIT PEPTIDE, 4FE-4S, 3D-STRUCTURE
477	1bab	B	2	140	6.8e-55			179.81	OXYGEN TRANSPORT HEMOGLOBIN THIONVILLE ALPHA CHAIN MUTANT WITH VAL 1 1BAB 3 REPLACED BY GLU AND AN ACETYLATED MET BOUND TO THE 1BAB 4 AMINO TERMINUS 1BAB 5	
477	1ch4	A	2	140	1.7e-55			168.27	MODULE-SUBSTITUTED CHIMERA HEMOGLOBIN BETA-ALPHA; CHAIN: A, B, C, D;	OXYGEN TRANSPORT OXYGEN TRANSPORT, CHIMERA PROTEIN, RESPIRATORY PROTEIN, HEME
477	1fdh	G	3	140	1e-55			150.09	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY, HUMAN FETAL F=/I1$=) 1FDHG 1 1FDHH 2	
477	1hda	B	3	140	8.5e-51			154.60	OXYGEN TRANSPORT HEMOGLOBIN (DEOXY) 1HDA 3	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
477	1ibc	B	2	140	1e-52			154.97	HEMOGLOBIN (DEOXY); CHAIN: A, B;	OXYGEN TRANSPORT HEME, OXYGEN TRANSPORT, RESPIRATORY PROTEIN, ERYTHROCYTE
477	1qpw	B	2	140	1e-52			163.36	PORICINE HEMOGLOBIN (ALPHA SUBUNIT); CHAIN: A, C; PORICINE HEMOGLOBIN (BETA SUBUNIT); CHAIN: B, D	OXYGEN TRANSPORT X-RAY STUDY, PORCINE HEMOGLOBIN, ARTIFICIAL HUMAN BLOOD, 2 OXYGEN TRANSPORT
480	1b6e		66	196	4.2e-29			81.88	CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
480	1bj3	A	67	193	3.4e-32			63.00	COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
480	1byf	A	77	194	5.1e-16			54.78	POLYANDROCARPA LECTIN; CHAIN: A, B;	SUGAR BINDING PROTEIN TC14; C-TYPE LECTIN, GALACTOSE-SPECIFIC, SUGAR BINDING PROTEIN
480	1esl		78	197	8.5e-31			53.21	CELL ADHESION PROTEIN E-SELECTIN (LECTIN AND EGF DOMAINS, RESIDUES 1 - 157) 1ESL 3 (FORMERLY KNOWN AS ELAM-1) 1ESL 4	
480	1htn		46	196	1e-26			58.22	TETRANECTIN; CHAIN: NULL;	LECTIN TETRANECTIN, PLASMINOGEN BINDING, KRINGLE 4, ALPHA-HELICAL 2 COILED COIL, C-TYPE LECTIN, CARBOHYDRATE RECOGNITION DOMAIN
480	1hup		46	194	1.7e-23			53.60	MANNOSE-BINDING PROTEIN; 1HUP 4 CHAIN: NULL; 1HUP 5	C-TYPE LECTIN ALPHA-HELICAL COILED-COIL 1HUP 12
480	1ixx	A	67	193	5.1e-30			60.13	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
480	1ixx	B	67	195	8.5e-32			69.01	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
										MOTIF, LOOP EXCHANGED DIMER
480	1lit		67	195	1.7e-33			77.94	LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
480	1qdd	A	51	195	6.8e-35			84.76	LITHOSTATHINE; CHAIN: A;	METAL BINDING PROTEIN PANCREATIC STONE PROTEIN. PSP; PANCREATIC STONE INHIBITOR. LITHOSTATHINE
480	1rtm	1	36	195	1e-22			50.50	LECTIN MANNOSE-BINDING PROTEIN A (CLOSTRIPAIN FRAGMENT) (CL-MBP-A) 1RTM 3 1RTM 96	
480	1tn3		62	196	5.1e-25			59.09	TETRANECTIN: CHAIN: NULL;	LECTIN TETRANECTIN, PLASMINOGEN BINDING, KRINGLE 4, C-TYPE LECTIN, 2 CARBOHYDRATE RECOGNITION DOMAIN
480	2msb	A	77	193	1.2e-21			53.42	LECTIN MANNOSE-BINDING PROTEIN A (LECTIN DOMAIN) COMPLEX WITH 2MSB 3 CALCIUM AND GLYCOPEPTIDE 2MSB 4	
489	1adq	L	21	235	3.4e-84			313.02	IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
489	1aqk	L	22	235	5.1e-83			285.37	FAB B7-15A2; CHAIN: L, H;	IMMUNOGLOBULIN HUMAN FAB, ANTI-TETANUS TOXOID, HIGH AFFINITY, CRYSTAL 2 PACKING MOTIF, PROGRAMMING PROPENSITY TO CRYSTALLIZE, 3 IMMUNOGLOBULIN
489	1bjm	A	20	235	6.8e-79			287.81	LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER; 1BJM 6 CHAIN: A, B; 1BJM 7	IMMUNOGLOBULIN BENCE-JONES PROTEIN; 1BJM 8 BENCE JONES, ANTIBODY, MULTIPLE QUATERNARY STRUCTURES 1BJM 13
489	1lil	A	21	235	1.2e-80			311.90	LAMBDA III BENCE JONES PROTEIN CLE: CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI BLAST	Verify Score	PMF Score	SeqFold Score	Compound	PDB Annotation
489	1mcw	W	20	235	8.5e-76			277.24	IMMUNOGLOBULIN IMMUNOGLOBULIN HETEROLOGOUS LIGHT CHAIN DIMER 1MCW 3 (/MCG$-/WEIR$ HYBRID) 1MCW 4	
489	1mfb	L	22	232	1.4e-96			225.27	IMMUNOGLOBULIN FAB FRAGMENT (MURINE SE155-4) COMPLEX WITH HEPTASACCHARIDE 1MFB 3 B: GAL(1-2)MAN(1-4)RAM(1-3)GAL(1-2)[ABE(1-3)]MAN(1-4)RAM 1MFB 4	
489	2fb4	L	22	235	8.5e-83			298.26	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB 2FB4 4	
489	2mcg	1	20	235	8.5e-81			292.66	IMMUNOGLOBULIN IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER (/MCG$) 2MCG 3 (TRIGONAL FORM) 2MCG 4	
489	7fab	L	20	231	1.4e-89			252.72	IMMUNOGLOBULIN IMMUNOGLOBULIN FAB' NEW (LAMBDA LIGHT CHAIN) 7FAB 3	
489	8fab	A	22	231	1.7e-81			313.64	IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	

TABLE 6

SEQ ID NO:	Position Of the Last Amino Acid Of Signal Peptide	Maximum Score	Mean Score
246	23	0.948	0.886
247	20	0.954	0.900
249	19	0.992	0.946
252	35	0.906	0.594
255	20	0.943	0.601
256	18	0.895	0.587
257	26	0.966	0.902
258	20	0.974	0.942
262	44	0.967	0.702
273	20	0.954	0.900
291	19	0.992	0.946
296	26	0.965	0.852
309	16	0.885	0.571
328	18	0.939	0.693
338	18	0.988	0.897
340	13	0.887	0.839
355	21	0.895	0.558
356	18	0.906	0.614
357	19	0.966	0.927
362	26	0.994	0.899
376	35	0.906	0.594
379	23	0.989	0.919
405	20	0.943	0.601
418	18	0.895	0.587
426	26	0.966	0.902
428	22	0.970	0.910
430	14	0.941	0.861
432	20	0.974	0.942
433	23	0.994	0.967
451	26	0.978	0.885
457	27	0.980	0.853
482	27	0.989	0.918
484	18	0.996	0.953
489	19	0.981	0.914

TABLE 7

SEQ ID NO:	Chromosomal location
1	6q27
2	4p16.3
3	4p16.3
4	1p21
5	8q13-q22
6	17
7	X
8	5
10	16
11	10
12	10
13	8pter-p23.3
15	17
16	X
17	11q23.2
18	19p13.3-p13.2
19	3p21.1
20	10
21	1
22	8
23	16
24	8
25	1
26	22q13.1
27	22q13.1
28	1
29	3
30	X
31	Xq27.3
32	Xq27.3
33	4
34	7q35-q36
35	11q12-1q22.2
36	11q23.1-q23.2
37	12
38	2q11.1-q11.2
39	17
40	7q32
41	22q13.2
42	1q42.13-q42.2
43	19q13.3
44	19p12
45	1q23.1-24.3
46	22q11.1-q11.2
48	17
49	8p22
50	22
51	3q23-q24
52	7p22-p21
53	16
54	12
55	21q22.3
56	18q
57	6
60	1
61	19
62	14

63	6q15-q16.1
64	13q12.3-q13.1
65	17q21-q22
66	7q11.2
67	12
68	12p13
69	19q13.13-q13.2
70	12
71	19
72	18
73	1p36.13-q31.3
74	14
75	7q21
76	7q21-q22
78	11p11.2-p11.1
80	22q13.31-q13.33
81	3p26-p25
82	2
84	22q13.2-q13.31
86	19
87	22q11.1-q11.2
88	17
89	7q11.21-q11.23
91	9
92	1p35.1-36.12
93	3q13.1-q13.2
94	15
95	19q13.2
96	1
97	20p11.1-11.22
98	19
100	6p12
101	3
102	3
103	X
104	3q29-qter
105	15
107	12
108	20p11.21-12.3
110	5
111	10
112	10
113	6p21.2-p21.3
114	12q15
115	22
118	19
119	Xp11.2
121	15
122	3
123	3
124	20
125	9
126	11q13
127	13
128	Xq21.1-Xq21.3
129	Xq28
130	19p13.1-p12
131	8q22-q23
133	17
134	1p36.3-p36.2

136	11p15.5
137	11p15.5
138	11p15.5
139	10p15-p13
140	3q29
141	11
142	20p12.2-13
143	20q13.3
144	19q13.3-q13.4
146	17
147	12p13.3
148	8q22
149	8q22
150	5
151	9q34
152	7q21
153	7p13-p12
154	Xp22.33
155	15
156	14
158	19q13.3
159	19q13.3
160	6
161	14q24.3
162	11
164	16
165	22q13.2-q13.31
166	19
167	11
168	5
169	1p34
170	8q11
171	8q11
172	17
173	19
176	19
177	11
178	7q22-q32
179	16q22.1
181	4q28
182	16p13.3
183	5
184	1
187	3p21.1-p14.3
188	17q21
189	7q21-q22
190	3p13-q26.1
191	17q21.2
192	3q27
193	22q13.2-13.3
194	11q22.2-q22.3
195	12q24.31-q24.32
196	19q13.4
197	17
198	17
199	16
200	20
201	20
202	5
203	17

204	11
205	20q11.2-q12
206	1q24-q41
207	17
208	14
209	11q13
210	6
211	17q21
212	6q21
214	16
216	17
217	6p21.31
219	Xp22
220	20
221	3
222	22q13.31-13.32
223	11q12
224	11q13.3
225	11q13.3
226	12
227	17q24-q25
228	20
229	9
230	11
231	15q24-q25
233	19q13.4
234	22q11.2
235	12p13
236	9
237	3p25-p24
238	14q24.3
240	19q13.3
241	20
242	6
243	16q21-q23
244	22q11.1-q11.2

TABLE 8

SEQ ID NO: of nucleotide sequence	SEQ ID NO: of polypeptide sequence	SEQ ID NO: in USSN 09/654,935 (Numbers to the right of the underscore correlate to sequence identifiers in USSN 09/654,935)
1	246	793_3
2	247	793_4
3	248	793_5
4	249	793_6
5	250	793_7
6	251	793_9
7	252	793_15
8	253	793_16
9	254	793_17
10	255	793_18
11	256	793_19
12	257	793_20
13	258	793_21
14	259	793_22
15	260	793_25
16	261	793_28
17	262	793_29
18	263	793_30
19	264	793_31
20	265	793_32
21	266	793_33
22	267	793_34
23	268	793_35
24	269	793_36
25	270	793_37
26	271	793_38
27	272	793_39
28	273	793_40
29	274	793_41
30	275	793_42
31	276	793_43
32	277	793_44
33	278	793_47
34	279	793_48
35	280	793_49
36	281	793_50
37	282	793_51
38	283	793_52
39	284	793_55
40	285	793_56
41	286	793_57
42	287	793_58
43	288	793_60
44	289	793_61
45	290	793_62
46	291	793_63
47	292	793_64
48	293	793_65
49	294	793_66
50	295	793_67
51	296	793_68
52	297	793_69

53	298	793_70
54	299	793_71
55	300	793_72
56	301	793_74
57	302	793_75
58	303	793_76
59	304	793_77
60	305	793_78
61	306	793_79
62	307	793_80
63	308	793_81
64	309	793_82
65	310	793_83
66	311	793_85
67	312	793_86
68	313	793_87
69	314	793_88
70	315	793_89
71	316	793_90
72	317	793_91
73	318	793_92
74	319	793_93
75	320	793_94
76	321	793_95
77	322	793_96
78	323	793_97
79	324	793_98
80	325	793_99
81	326	793_101
82	327	793_102
83	328	793_103
84	329	793_104
85	330	793_106
86	331	793_107
87	332	793_108
88	333	793_109
89	334	793_110
90	335	793_111
91	336	793_112
92	337	793_113
93	338	793_114
94	339	793_115
95	340	793_116
96	341	793_117
97	342	793_118
98	343	793_119
99	344	793_120
100	345	793_121
101	346	793_122
102	347	793_123
103	348	793_124
104	349	793_125
105	350	793_126
106	351	793_127
107	352	793_128
108	353	793_129
109	354	793_130
110	355	793_131
111	356	793_132
112	357	793_133

113	358	793_134
114	359	793_135
115	360	793_136
116	361	793_137
117	362	793_138
118	363	793_139
119	364	793_140
120	365	793_141
121	366	793_142
122	367	793_143
123	368	793_144
124	369	793_145
125	370	793_146
126	371	793_147
127	372	793_148
128	373	793_149
129	374	793_150
130	375	793_151
131	376	793_152
132	377	793_153
133	378	793_154
134	379	793_155
135	380	793_156
136	381	793_157
137	382	793_158
138	383	793_159
139	384	793_160
140	385	793_161
141	386	793_162
142	387	793_163
143	388	793_164
144	389	793_165
145	390	793_166
146	391	793_167
147	392	793_168
148	393	793_169
149	394	793_170
150	395	793_171
151	396	793_172
152	397	793_173
153	398	793_174
154	399	793_175
155	400	793_176
156	401	793_177
157	402	793_178
158	403	793_179
159	404	793_180
160	405	793_181
161	406	793_182
162	407	793_183
163	408	793_184
164	409	793_185
165	410	793_186
166	411	793_187
167	412	793_188
168	413	793_189
169	414	793_190
170	415	793_191
171	416	793_192
172	417	793_193

173	418	793_194
174	419	793_195
175	420	793_196
176	421	793_197
177	422	793_198
178	423	793_200
179	424	793_201
180	425	793_202
181	426	793_203
182	427	793_204
183	428	793_205
184	429	793_206
185	430	793_207
186	431	793_209
187	432	793_210
188	433	793_211
189	434	793_212
190	435	793_213
191	436	793_214
192	437	793_215
193	438	793_216
194	439	793_217
195	440	793_218
196	441	793_219
197	442	793_220
198	443	793_221
199	444	793_222
200	445	793_223
201	446	793_224
202	447	793_225
203	448	793_226
204	449	793_227
205	450	793_229
206	451	793_230
207	452	793_231
208	453	793_232
209	454	793_233
210	455	793_234
211	456	793_235
212	457	793_236
213	458	793_237
214	459	793_238
215	460	793_239
216	461	793_240
217	462	793_241
218	463	793_242
219	464	793_244
220	465	793_245
221	466	793_247
222	467	793_248
223	468	793_249
224	469	793_250
225	470	793_251
226	471	793_252
227	472	793_253
228	473	793_254
229	474	793_255
230	475	793_256
231	476	793_257
232	477	793_258

233	478	793_259
234	479	793_260
235	480	793_261
236	481	793_262
237	482	793_263
238	483	793_264
239	484	793_265
240	485	793_266
241	486	793_267
242	487	793_268
243	488	793_269
244	489	793_270
245	490	793_271

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected from SEQ ID NO: 1-245, a mature protein coding portion of SEQ ID NO: 1-245, an active domain coding portion of SEQ ID NO: 1-245, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-245, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides SEQ ID NO: 2146-490, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-245.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 March 2002 (07.03.2002)

PCT

(10) International Publication Number
WO 02/018424 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, 16/18, G01N 33/53, 33/50, C12N 5/10, C12Q 1/68

(21) International Application Number: PCT/US01/27093

(22) International Filing Date: 31 August 2001 (31.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/654,935 1 September 2000 (01.09.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/654,935 (CIP)
Filed on 1 September 2000 (01.09.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHOU, Ping** [US/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendahl Lane, Cupertino, CA 95014 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **WANG, Dunrui** [CN/US]; 12252 Pepper Tree Lane, Poway, CA 92064 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [US/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). **WEHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P., C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(88) Date of publication of the international search report:
15 May 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/018424 A3

(54) Title: NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/27093

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 G01N33/53 G01N33/50
C12N5/10 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, EMBL, BIOSIS, MEDLINE, PAJ, WPI Data, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 01 53312 A (CHEN RUI HONG ;GOODRICH RYLE (US); HYSEQ INC (US); WANG DUNRUI (US) 26 July 2001 (2001-07-26) SEQ ID NO:4445, 6231 ---	1-28
X	DATABASE EMBL [Online] 19 January 1998 (1998-01-19) PHILIPPS, S:: "Human DNA sequence from clone 366N23 on chromosome 6q27. Contains two genes similar to consecutive parts of the C. elegans UNC-93 (protein 1, C46F11.1) gene, a KIAA0173 and Tubulin-Tyrosine Ligase LIKE gene, a Mitotic Feedback..." retrieved from EBI Database accession no. AL021331 XP002214453 abstract --- -/--	1-28

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 September 2002	14 01 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Schmitz, T

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE SWALL [Online] 1 November 1998 (1998-11-01) PHILLIPS, S.: "DJ366N23.1 (Putative C. elegans UNC-93 (Protein 1, C46F11.1) like protein)." retrieved from EBI Database accession no. 075651 XP002214454 abstract --- DATABASE EMBL [Online] 18 July 1996 (1996-07-18) HILLIER L. ET AL.: "zh48g06.r1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:415354 5' similar to PIR:S23352 S23352 gene unc-93 protein 1 - Caenorhabditis elegans [1] ;, mRNA sequence." retrieved from EBI Database accession no. W92071 XP002214455 abstract -----	1-28
X		10-12, 16-20, 27,28

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/27093

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 27, 28 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 13, 16 (partially)
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-28 all partially

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 13, 16 (partially)

Present claims 13, 16 relate to a compound defined by reference to a desirable characteristic or property, namely the binding to the claimed polypeptides.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the antibody binding to the claimed polypeptide.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1: 1-28 (all partially)

SEQ ID NO:1, 246.
Furthermore vectors, host cells, methods, collections, antibodies, all referring to said nucleotide or amino acid sequence.

Invention 2: 1-28 (all partially)

As invention 1, but referring to SEQ ID NO:2, 247.

Invention 3: 1-28 (all partially)

As invention 1, but referring to SEQ ID NO:3, 248.

.
.
.

Invention 245: 1-28 (all partially)

As invention 1, but referring to SEQ ID NO:245, 490.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/27093

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0153312 A	26-07-2001	AU 2292401 A	31-07-2001
		AU 2591801 A	31-07-2001
		AU 2593601 A	31-07-2001
		AU 2595501 A	31-07-2001
		AU 2596501 A	31-07-2001
		AU 2598301 A	31-07-2001
		AU 2728401 A	31-07-2001
		AU 2734401 A	31-07-2001
		AU 2734801 A	31-07-2001
		AU 2738501 A	31-07-2001
		AU 3265701 A	31-07-2001
		EP 1242596 A1	25-09-2002
		EP 1240178 A2	18-09-2002
		EP 1242580 A1	25-09-2002
		EP 1242443 A1	25-09-2002
		EP 1250346 A2	23-10-2002
		EP 1254256 A2	06-11-2002
		EP 1248848 A1	16-10-2002
		WO 0153312 A1	26-07-2001
		WO 0153453 A2	26-07-2001
		WO 0153326 A1	26-07-2001
		WO 0153454 A2	26-07-2001
		WO 0153455 A2	26-07-2001
		WO 0153456 A2	26-07-2001
		WO 0153466 A1	26-07-2001
		WO 0152616 A2	26-07-2001
		WO 0153500 A1	26-07-2001
		WO 0153515 A1	26-07-2001
		WO 0153485 A1	26-07-2001
		US 6465620 B1	15-10-2002
		US 2002146692 A1	10-10-2002
		AU 5362001 A	30-10-2001
		WO 0179446 A2	25-10-2001
		US 2002142953 A1	03-10-2002